



# The COP9 signalosome: Activity and regulation

Melissa Birol

## ► To cite this version:

Melissa Birol. The COP9 signalosome: Activity and regulation. Biochemistry, Molecular Biology. Université Montpellier II - Sciences et Techniques du Languedoc, 2014. English. NNT: 2014MON20072 . tel-01386906

**HAL Id: tel-01386906**

**<https://theses.hal.science/tel-01386906>**

Submitted on 24 Oct 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE

Pour obtenir le grade de  
Docteur

Délivré par **UNIVERSITE MONTPELLIER 2**

Préparée au sein de l'école doctorale Sciences Chimiques  
et Biologiques pour la Santé

Et de l'unité de recherche UMR 5048 - UMR 5048 - Centre  
de Biochimie Structurale - CBS

Spécialité : Biologie Santé

Présentée par **Melissa Birol**

**Le COP9 signalosome: Activité et régulation**  
**The COP9 signalosome: Activity and  
regulation**

Soutenue le 19/12/2014 devant le jury composé de

Mme Aude Echalié, MC, Department of Biochemistry, University of Leicester	Directeur de thèse
Mr Christian Dumas, DR, Centre de biochimie Structurale	Directeur de these
Mme Elisabetta Bianchi, MD, Institut Pasteur	Rapporteur
Mr Dominique Housset, Ingénieur-chercheur CEA, Institut de Biologie Structurale	Rapporteur
Mr Alain Chavanieu, DR MC - UM1, Faculté de Pharmacie	Examineur
Mme Thérèse Malliavin, CR1 CNRS, Institut Pasteur	Examineur
Mr Marc Piechaczyk, DR CNRS, L'institut de Génétique Moléculaire de Montpellier	Examineur



# ***The COP9 signalosome: Activity and regulation***

**By Melissa Birol**

Université Montpellier 2  
Centre de Biochimie Structurale  
December 2014



## Acknowledgments

I would first and foremost like to express my gratitude to my supervisors Dr Aude Echali r and Dr. Christian Dumas for their endless knowledge and creativity throughout this project. With their broad personal experience they guided me toward accurate and truthful scientific research of high standards. I would like to thank them for their indisputable support, availability and most of all his friendship in all times. I have enjoyed science with them. Dr. Aude Echali r has been my mentor/and a very good friend, and will be continue to be. I would not be able to find words to express how grateful I am to her for the amount to knowledge she has transferred to me in the areas of biochemistry and structural biology and for engaging me in such a challenging but yet exciting project. Dr. Christian Dumas has most evidently showed me how exciting science can be. Conversations with him were overwhelming with the amount of knowledge that spilled. I thank him for all the insight and help he provided me throughout these three years.

Many thanks to my PhD committee members, Dr. O. Coux and Dr. D. Xirodimas for their support, encouragement, and insight over the three years. Their comments and advice has been extremely helpful for the direction this project has taken.

Many thanks to collaborators who have participated in the work performed. I would like to acknowledge Dr. D. Xirodimas for sharing his knowledge and bring insight in cellular biology to the project and also in his guidance in the *in vivo* work performed on the work described in Chapter 5. This includes, his useful advises, his help with cell growth and encouraging conversations in the final part of this project. I would also like to thank F. Hoh for his technical help in X-ray crystallography. Additional thanks to Dr. A. Padilla and X. Yang for their participation in NMR data collection and analysis. Also would like to acknowledge the contribution of C. Ebel, F.X. Clare, Prof. M. Peter and Dr. R. I. Enchev in the work presented, both technically as well as in the scientific discussions and reagent preparations.

I gratefully acknowledge the funding sources that made my Ph.D. work possible. I was funded by the Ph. D. fellowship awarded at the Universit  Montpellier 2 by the Doctoral school CBS2.

I would like to thank everyone at Centre de Biochimie Structurale (CBS), past and present, for making the lab a place worth to come to every day: Special thanks are deserved by my friends and colleagues, Selma, Laurent, Desiree, Hussein, Anais, Antoine, for infinite scientific debates that intrigued our mind and for their true friendship. Memories of these three years are engraved and I take them with me to remember just what great friends I have got. Additionally other Ph.D. and Postdocs within the institute have always been there for insightful conversations. I thank you for your insights and advise along the way. I will take it with me where ever I'll go. Would like to thank friends outside my work environment who have encouraged me

with my PhD including, Kallina, Kate, Maria-Alexia, Alki, Ioana, Alex, George, Maike, Marco, and Nikos, for always being there for me. To reach this stage I have passed through various steps, all of which have allowed me to make friendship with various people who are not named here but nevertheless have marked a special place.

Above all I would like to thank my family, also including Iliana, Basili and Christopher, for their support throughout the years. I thank my parents for believing in me and leading me the way and giving me the opportunity to pursue education and providing me with the tools for becoming the daughter that they deserve to have. I lastly want to thank my Grandparents for inspiring me throughout.

## Table of Content

<b>List of Figures and Tables .....</b>	<b>6</b>
Figures.....	6
Tables.....	6
<b>Abbreviations</b>	
<b>Résumé court .....</b>	<b>10</b>
<b>Résumé long .....</b>	<b>11</b>
1. Introduction .....	11
2. Résultats .....	13
3. Conclusion.....	18
<b>Abstract.....</b>	<b>21</b>
<b>Chapter 1. INTRODUCTION.....</b>	<b>22</b>
<b>1.1 Biology:</b>	
1.1.1 The ubiquitin-proteasome system .....	22
1.1.2 E3 ligases.....	23
1.1.3 Cullin-RING Ub ligases.....	25
1.1.4 Proteasomal degradation.....	26
1.1.5 Nedd8 - An Ub-like protein .....	27
1.1.6 The CSN .....	29
1.1.7 SCF regulation by the CSN complex.....	30
<b>1.2 Cancer:</b>	
1.2.1 The CSN is overexpressed in cancers .....	33
1.2.2 The p27 axis.....	34
1.2.3 The p53 axis.....	35
1.2.4 The p57 axis.....	36
<b>1.3 CSN in different species:</b>	
1.3.1 The architecture of the CSN in different species.....	36
1.3.2 The mammalian CSN complex .....	37
1.3.3 The CSN in plants.....	38
1.3.4 CSN in yeast and fungi.....	39
<b>1.4 CSN isopeptidase and other isopeptidases:</b>	
1.4.1 Isopeptidases.....	40
1.4.2 The CSN Isopeptidase.....	42
1.4.3 The 19S proteasome lid isopeptidase.....	43
1.4.4 The eIF3 isopeptidase .....	44
1.4.5 The BRCC36-containing isopeptidases .....	44
1.4.6 The AMSH-LP isopeptidase.....	45
<b>Chapter 2: Materials and Protocols.....</b>	<b>48</b>
<b>2.1 Protein production:</b>	
2.1.1 CSN5 Construct design, cloning, expression and purification .....	48
2.1.2 CSN6 Construct design, cloning, expression and purification .....	48
2.1.3 Construct design, cloning, expression and purification of other proteins used in this project.....	49
<b>2.2 Enzymatic assays:</b>	
2.2.1 Isopeptidase assays on synthetic substrates.....	49
2.2.2 Pro-Nedd8 processing assays .....	50

2.2.3 Deneddylation assays .....	50
<b>2.3 Biophysical characterisation:</b>	
2.2.1 Fluorescence anisotropy .....	51
2.2.2 Isothermal titration calorimetry .....	51
2.3.3 Crystallisation, data collection and structure determination of human CSN6. ....	52
2.3.4 Approaches used for CSN5-CSN6 crystallisation trials. ....	52
2.3.5 Molecular dynamics simulations. ....	53
2.3.6 NMR. ....	54
2.3.7 Protein-protein docking and modelling. ....	54
2.3.8 Analytical ultracentrifugation (Performed by Dr. C. Ebel, IBS, Grenoble) .....	55
<b>2.4 In vivo approaches:</b>	
2.4.1 Co-immunoprecipitation experiments (Performed by Dr. F.-X. Claret, MD Anderson, Houston, USA) .....	55
2.4.2 Growth of NEDP1 MEFs cells (Performed by Dr. D. Xirodimas, CRBM, Montpellier) ...	56
2.4.3 Enzymatic assays performed on cell lysates .....	56
<b>Chapter 3: CSN5 isopeptidase activity .....</b>	<b>57</b>
<b>3.1 Introduction:</b>	
<b>3.2 Structure of CSN5:</b>	
3.2.1 CSN5 .....	65
3.2.1.1 CSN5 1-257 crystal structure .....	65
3.2.1.2 Model and structure of CSN5 C-terminal extension .....	69
3.2.2 A structural comparison between AMSH-LP, AMSH, and Rpn11 .....	70
3.2.2.1 MPN domain .....	70
3.2.2.2 Molecular models and analyses of other MPN+/JAMM proteins.....	72
3.2.2.3 Structural elements: Zinc-binding site, Ins-1, Ins-2 .....	73
3.2.3 Molecular basis of CSN5 auto-inhibition .....	78
3.2.3.1 Implication of AMSH-LP and of AMSH Ins-1 in substrate recruitment.....	78
3.2.3.2 Auto-inhibition of CSN5 through the Ins-1 conformation .....	78
3.2.4 Dynamics of CSN5 .....	80
3.2.4.1 Regions with high flexibility signs .....	81
3.3.3 Unlocking the active site lock of CSN5 .....	84
<b>3.4 Activation of an auto-inhibited CSN5</b>	
3.4.1 Isopeptidase activity of CSN5.....	86
3.4.2 Deneddylation activity of CSN5 .....	87
3.4.3 Molecular basis of CSN5 inactive state.....	89
3.4.3.1 Nedd8 binding.....	89
<b>3.5 MPN homo-dimerisation</b>	
3.5.1 CSN5 homodimerisation in crystal structure.....	92
3.5.2 CSN5 homodimerisation in solution and <i>in vivo</i> .....	95
3.5.3 Study of CSN5 homodimerisation: evaluation of the most probable CSN5 dimer in cell extracts.....	97
3.5.3.1 The dynamics of CSN5 dimers .....	98
3.5.3.2 Validation of CSN5 dimer interface .....	99
<b>Chapter 4: CSN5 activity within the CSN complex.....</b>	<b>102</b>
<b>4.1 Introduction:</b>	
<b>4.2 CSN5 directly binds to CSN6:</b>	
4.2.1 Investigation of the direct interaction between CSN5 and CSN6 .....	104
4.2.2 Affinity of the complex .....	104
<b>4.3 CSN6 stimulates CSN5 isopeptidase activity:</b>	
4.3.1 CSN5/CSN6 activity on synthetic substrates.....	106
4.3.2 The CSN5/CSN6 complex recruits Nedd8.....	111
4.3.3 Cullin deneddylation by CSN5-CSN6 complex .....	112
<b>4.4 Human CSN6 MPN adopts a structure that is reminiscent of <i>Hs</i> Rpn8:</b>	

4.4.1 Crystal Structure of CSN6 .....	115
4.4.2 CSN6 conservation and molecular surface properties.....	119
4.4.3 Comparison of CSN5, an MPN+/JAMM protein with structurally characterised MPN- proteins .....	119
<b>4.5 Structural characterisation of the CSN5/CSN6 complex:</b>	
4.5.1 Crystallisation of the heterodimer complex .....	123
4.5.2 Mapping the interface of the CSN5/CSN6 complex .....	125
4.5.3 Interfacial mutations on the CSN5-CSN6 complex.....	128
<b>4.6 Construction of the ternary CSN5/CSN6/Nedd8 complex:</b>	
4.6.1 Investigation of Nedd8 and CSN5 interaction.....	133
4.6.2 Model of CSN5-Nedd8 complex.....	136
4.6.3 Model of the ternary complex, CSN5-CSN6-Nedd8.....	136
<b>Chapter 5: Conclusions and Discussion.....</b>	<b>140</b>
<b>Annex.....</b>	<b>162</b>
<b>Paper 1</b>	
Insights into the regulation of the human COP9 signalosome catalytic subunit, CSN5/Jab1. .....	162
<b>Paper 2</b>	
Structure and Function of MPN (Mpr1/Pad1 N-Terminal) Domain-Containing Proteins ..	163
<b>Paper 3</b>	
Structural and Biochemical Characterization of the COP9 Signalosome CSN5/CSN6 Heterodimer.....	164
<b>Patent filed</b>	
Csn5 polypeptides and uses thereof for screening therapeutic agents .....	165

# List of Figures and Tables

## Figures

FIGURE 1-1. SCHEMATIC OVERVIEW OF THE UPS. ....	23
FIGURE 1-2. E3 LIGASES. ....	24
FIGURE 1-3. STRUCTURE OF Ub AND UBL PROTEINS. ....	28
FIGURE 1-4. CRL FUNCTIONAL CYCLE. ....	32
FIGURE 1-5. THE HUMAN DUBOMES. ....	41
FIGURE 1-6. SCHEMATIC REPRESENTATION OF MPN+/MPN- DOMAIN-CONTAINING MULTI-PROTEIN COMPLEXES. ....	46
FIGURE 3-1 SCHEMATIC REPRESENTATION OF THE ACTIVE SITE OF CYSTEINE-DEPENDENT PROTEASES SHOWING THE MECHANISM OF CLEAVAGE OF AN ISOPEPTIDE BIND. ....	58
FIGURE 3-2. STRUCTURE OF CSN5 MPN DOMAIN. ....	66
FIGURE 3-3. COMPARISON OF AMSH-LP AND CSN5 INS-1 REGION. ....	67
FIGURE 3-4. N- AND C-TERMINAL EXTENSIONS OF CSN5. ....	69
FIGURE 3-5. THE JAMM/MPN+ MEMBERS. ....	73
FIGURE 3-6. THE CATALYTIC MPN CORE DOMAIN. ....	77
FIGURE 3-7. CSN5 CONFORMATION IN THE STAND-ALONE FORM IS NOT COMPATIBLE WITH UBL BINDING. ....	79
FIGURE 3-8. CSN5 MPN DOMAIN AND SPECIFIC EXTENSIONS. ....	83
FIGURE 3-9. RELAXATION OF THE CSN5 INS-1 CONFORMATION. ....	85
FIGURE 3-10. ISOPEPTIDASE ACTIVITY OF CSN5. ....	90
FIGURE 3-11. CSN5 OLIGOMERIC STATE. ....	94
FIGURE 3-12. OLIGOMERIC ARRANGEMENT OF CSN5. ....	97
FIGURE 3-13. DYNAMICS OF CSN5 HOMODIMERS. ....	99
FIGURE 3-14. DISRUPTING THE OLIGOMERIC STATE OF CSN5. ....	101
FIGURE 4-1. CSN5 AND CSN6 FORM A BINARY COMPLEX. ....	105
FIGURE 4-2. CSN6 ACTIVATES CSN5 TO FORM THE CATALYTIC CORE OF THE CSN COMPLEX. ....	110
FIGURE 4-3. DETERMINATION OF THE DISSOCIATION CONSTANT OF THE CSN5-CSN6 INTERACTION WITH NEDD8. ....	111
FIGURE 4-4. DENEDDYLASE ACTIVITY OF THE CSN5-CSN6 HETERODIMER. ....	114
FIGURE 4-5. CRYSTAL STRUCTURE OF HUMAN CSN6. ....	117
FIGURE 4-6. HOMODIMERISATION OF MPN- DOMAIN CONTAINING PROTEINS. ....	118
FIGURE 4-7. BIOCHEMICAL AND STRUCTURAL PROPERTIES OF HUMAN CSN6. ....	120
FIGURE 4-8. ELUCIDATION OF THE INTERFACE BETWEEN THE CSN5-CSN6 HETERODIMER. ....	127
FIGURE 4-9. MUTATIONS PERFORMED TO VALIDATE THE CSN5-CSN6 HETERODIMER INTERFACE. ....	130
FIGURE 4-10. NEDD8 BINDING TO CSN5. ....	134
FIGURE 4-11. JAMM/MPN+-MPN- HETERODIMER ASSEMBLY. ....	138
FIGURE 5-1. MODES OF ACTIVITY REGULATION OF CSN5 AND THE CSN COMPLEX. ....	144

## Tables

TABLE 1-1. CSN AND 19S PROTEASOME LID SUBUNIT COMPOSITION IN DIFFERENT ORGANISMS. ....	37
TABLE 3-1. RMSD VALUES ACROSS THE MPN SUB-FAMILY AGAINST CSN5 (PDB CODE: 4F7O). ....	70
TABLE 3-2. DISTANCES BETWEEN ZN AND LIGAND ATOMS IN THE ACTIVE SITE OF THE CSN5 JAMM MOTIF. ....	75
TABLE 3-3. CSN5 INTERFACE SURFACE PROPERTIES OBSERVED IN THE CRYSTAL STRUCTURE. ....	93
TABLE 4-1. DATA COLLECTION AND STRUCTURE REFINEMENT STATISTICS OF HUMAN CSN6 MPN DOMAIN. ....	116
TABLE 4-2 SUMMARY OF THE MPN-MPN CONTACTS (IN THE CONTEXT OF HOMO- OR HETERO-DIMERS) IN MPN CRYSTAL STRUCTURES. ....	122
TABLE 4-3. CSN6 RESIDUES FOUND UNPERTURBED, UPON BINDING WITH CSN5 AS EVALUATED BY NMR HSQC. ....	127

## Abbreviations

AMC: 7-amino-4-methyl-coumarin

AMSH: associated molecule with the SH3 domain of STAM

AMSH-LP: AMSH-like protein

*An: Aspergillus nidulans*

*At: Arabidopsis thaliana*

BRCA1/2: Breast Cancer 1/2

BRCC36: BRCA1/BRCA2-Containing Complex subunit 36

BRCC45: BRCA1/BRCA2-containing complex subunit 45

BS3: bis-(sulfosuccinimidyl)suberate

BSA: bovine serum albumine

CAND1: Cullin-associated NEDD8-dissociated protein 1

CDK: Cyclin-dependent kinase

CDKN1C: Cyclin-dependent kinase inhibitor 1C

*Ce: Caenorhabditis elegans*

CRL: Cullin RING E3 ubiquitin ligase

CSN: COP9 signalosome

COP9: COntstitutive Photomorphogenesis 9

Cul1: Cullin 1

CX: cross-linking

CYLD: cylindromatosis gene product

*Dm: Drosophila melanogaster*

DSS: disuccinimidyl suberate

DUB: deubiquitinating enzyme

eIF3: eukaryotic translation initiation factor 3

EM: electron microscopy

HECT: homologous to E6-associated carboxyl terminus

HIF-1 $\alpha$ : hypoxia-inducible factor 1- $\alpha$

HSQC: Heteronuclear single quantum coherence spectroscopy

*Hs: Homo sapiens*

hVIP: human immunodeficiency virus 1 accessory protein

Id1: Inhibitor of DNA binding 1  
 ITC: isothermal titration calorimetry  
 Jab1: Jun-activatory binding protein 1  
 JAMM: Jab1-MPN-Mov34  
 $K_D$ : dissociation constant  
 KO: knockout  
 Mdm2: Double minute 2 protein  
 MEF: mouse embryonic fibroblasts  
 MERIT40: Mediator of RAP80 interactions and targeting subunit of 40 kDa  
*Mm: Mus musculus*  
 MPN: Mpr1/Pad1 N-terminal  
 NAE: Nedd8-specific activating enzyme  
*Nc: Neurospora crassa*  
 Nedd8: neural precursor cell expressed developmentally down-regulated protein 8  
 NEDP1: Nedd8-specific protease 1  
 NMR: nuclear magnetic resonance  
 OUT: ovarian-tumor  
 PCI: Proteasome COP9 eIF3  
 Pro-Nedd8: precursor Nedd8  
 RAP80: Receptor-associated protein 80  
 RING: really interesting new gene  
 RBR: RING-between-RING  
 Runx3: Runt-related transcription factor 3  
 Rbx1: RING-box protein 1  
*Sc: Saccharomyces cerevisiae*  
 SUMO: Small Ubiquitin-related MOdifier  
 Skp1: S-phase kinase-associated protein 1  
 Skp2: S-phase kinase-associated protein 2  
 SCF: Skp1-Cullin1-F-box  
 Smad7: Mothers against decapentaplegic homolog 7  
*Sp: Schizosaccharomyces pombe*  
 STAM1: signal transduction adaptor molecule 1  
 Ub: ubiquitin



UBL: ubiquitin-like molecule

UCH: Ub C-terminal hydrolases

UPS: ubiquitin proteasome system

USP: Ub-specific protease

WH: winged-helix

$\beta$ -TrCP:  $\beta$ -transducin repeat-containing protein

## Résumé court

Le COP9 (Constitutive photomorphogenesis 9) signalosome (CSN) est un complexe multiprotéique contenant huit sous-unités (320 kDa), impliqué dans des processus cellulaires divers allant de la progression du cycle cellulaire, à l'expression des gènes et la réparation de l'ADN, à travers sa fonction au sein du système ubiquitine-protéasome. Il s'agit d'un complexe fortement conservé au cours de l'évolution chez les eucaryotes supérieurs chez qui son activité catalytique est essentielle. Au cours des années d'études biologiques et biochimiques qui ont permis d'élucider le rôle du CSN, sa fonction la mieux étudiée et la mieux comprise est celle liée au contrôle de l'ubiquitylation des protéines (une modification post-traductionnelle qui implique la liaison covalente d'une protéine cible par une molécule d'ubiquitine) par une classe d'E3 ubiquitine ligases. L'activité catalytique du CSN régule spécifiquement les E3 culline RING ubiquitin ligases (CRLs) via la suppression d'une molécule ressemblant à l'ubiquitine, Nedd8 (cullin-neural precursor cell expressed developmentally downregulated gene 8) des CRLs au cours d'une réaction de déneddylation. Les cycles de neddylation/dénéddylation sont essentiels au fonctionnement correct des CRLs et le CSN joue un rôle central dans ce processus à travers sa fonction de déneddylase. Une similarité globale lie le CSN, le chapeau du protéasome (19S) et le complexe eIF3 (eukaryotic initiation factor-3). Ces assemblages multi-protéiques comprennent tous six sous-unités contenant un domaine PCI (proteasome COP9 eIF3) et deux sous-unités contenant un domaine MPN (Mpr1–Pad1–N-terminal). L'activité catalytique du CSN est portée par la sous-unité 5, CSN5, qui hydrolyse la liaison isopeptidique entre Nedd8 et la CRL. CSN5 contient un cœur catalytique dépendant d'un zinc et comprenant un motif JAMM (Jab1/MPN/Mov34).

L'incorporation de CSN5 dans le CSN révèle son activité isopeptidasique, alors qu'à l'état isolé, CSN5 n'est pas actif. Le travail réalisé au cours de ces trois ans a abouti à quatre aspects principaux qui ont contribué à une meilleure compréhension globale du système CSN. (i) S'appuyant sur la structure du domaine catalytique de CSN5, des études *in vitro* et *in silico* ont abouti à l'identification d'un élément moléculaire permettant à CSN5 de passer de la forme inactive à la forme active. Ceci a débouché sur la conception et validation d'un variant constitutivement actif de CSN5. (ii) La capacité de CSN5 à homodimériser a été étudiée en solution, *in silico* et dans des extraits cellulaires et a apporté des perspectives potentiellement intéressantes concernant la fonction de CSN5. (iii) Au delà de ce travail et pour aborder la question de la régulation de l'activité de CSN5 dans le CSN, la contribution de la sous-unité 6, CSN6 qui interagit directement avec CSN5 a été évaluée et ceci a abouti à l'identification de CSN6 comme sous-unité activatrice de CSN5. (iv) La caractérisation biochimique et biophysique du complexe CSN5-CSN6 a été utilisée pour explorer les bases moléculaires de cette association, non seulement, dans le contexte de son interaction avec Nedd8, mais aussi, de son intégration au sein du CSN, à travers une approche intégrée alliant des techniques biochimiques, structurales, biophysiques et computationnelles.

En résumé, ce travail a permis d'améliorer la compréhension des déterminants de l'activité et des mécanismes de régulation auxquels la sous-unité catalytique du CSN, CSN5 est soumise.

**Mots clés:** système ubiquitine-protéasome ; Cop9 signalosome; activité isopeptidasique; ubiquitin E3 ligase; régulation de Cullin, Rpn11, pro-Nedd8

## Résumé long

### 1. Introduction

---

L'ubiquitination des protéines est une modification post-traductionnelle qui est au cœur du système ubiquitine-protéasome et qui intervient dans la quasi-intégralité des voies de signalisation cellulaires chez les eucaryotes. La plus connue des conséquences de l'ubiquitination est l'adressage des protéines marquées vers la machinerie de dégradation 26S protéasome, mais l'ubiquitination des protéines peut induire d'autres effets, comme la modification de l'interactome de la protéine modifiée, de son activité. Le contrôle de l'ensemble de ces processus liés au système ubiquitine-protéasome est un aspect central à l'homéostasie cellulaire. L'ubiquitination d'une molécule fait intervenir trois réactions en cascade (E1, E2, E3). La fonction carboxylique de la glycine d'une ubiquitine est d'abord activée par l'enzyme E1 (ubiquitin activating enzyme). La seconde étape comprend le transfert de l'ubiquitine activée de l'E1 vers la cystéine catalytique de l'enzyme E2 (ubiquitin conjugating enzyme). L'enzyme E2 chargée avec l'ubiquitine s'associe à l'E3 en interaction avec le substrat et ensemble l'E2 et l'E3 coopèrent pour le transfert de l'ubiquitine sur la chaîne latérale d'une lysine de surface de la protéine cible, formant une liaison isopeptidique. C'est le dernier protagoniste de cette réaction multi-étape de transfert des ubiquitines sur les protéines cibles, les E3 ligases qui nous intéressent tout particulièrement dans le contexte de ce travail.

Il existe plusieurs classes d'E3 ligases définies selon leur composition protéique. Les E3 Culline RING ubiquitine ligases (CRLs ; RING for really interesting new gene) constituent une des classes majeures d'E3 ubiquitine ligases. Ces CRLs comprennent plusieurs sous-unités, parmi lesquelles la protéine d'échafaudage, culline et la sous-unité RING porteuse de l'activité enzymatique. L'activité de ces CRLs est régulée par des cycles de neddylation/déneddylation (attachement covalent d'une molécule proche de l'ubiquitine, Nedd8) sur la sous-unité culline.

Les CRLs possèdent une architecture complexe et dynamique qui reflète les différents mécanismes de régulation dont elles sont l'objet. Parmi ces voies de contrôle, la neddylation des CRLs au niveau de la sous-unité d'échafaudage, Culline, stimule l'activité d'ubiquitylation des CRLs à travers de larges changements conformationnels de leur structure globale et empêche la fixation de l'inhibiteur naturel des CRLs, CAND1 (Cullin-associated NEDD8-dissociated protein 1). Les cycles de neddylation/déneddylation sont strictement régulés et un élément central de cette régulation est le complexe COP9 signalosome (CSN) qui déneddyle les Cullines (c'est à dire hydrolyse la liaison isopeptidique qui lie Nedd8 à la Culline). En effet, il a été montré que le CSN joue un rôle critique dans le contrôle de l'état de neddylation des CRLs. Au delà de l'activité catalytique de déneddylation du CSN sur les CRLs, CSN peut aussi interagir de manière non-catalytique avec les CRLs. Ce mode de régulation non-catalytique, encore mal compris, pourrait augmenter encore la complexité de

ce système de régulation liant le CSN aux CRLs qui échappe toujours à une compréhension parfaite. Bien que le CSN soit *in vitro* un inhibiteur des CRLs, il a été montré que l'activité du CSN est essentielle aux CRLs *in cellulo*. Plusieurs mécanismes, non mutuellement exclusifs, permettant de comprendre ce paradoxe ont été proposés. (i) Le CSN pourrait par son activité déneddylase empêcher l'auto-ubiquitination des CRLs qui conduit à leur dégradation. (ii) La nécessité non seulement de la liaison covalente de Nedd8 mais aussi de son hydrolyse suggère qu'*in vivo* les CRLs sont sujettes à des cycles rapides de neddylation/dénéddylation. Un cycle de neddylation/dénéddylation pourrait correspondre à un événement de fixation du substrat, événement lié à la neddylation de la Culline auquel ferait suite l'ubiquitination du substrat, sa dissociation et la déneddylation de la CRL. A la fin de ce cycle, un nouveau cycle serait entamé par la fixation du substrat. (iii) Le CSN joue aussi un rôle dans la régulation de la fixation d'un inhibiteur naturel des CRLs, CAND1. CAND1 interagit spécifiquement avec les sous-unités Cullines qui sont libérées de leur modification par Nedd8 et de la sous-unité de recrutement du substrat. Le CSN pourrait, dans ce cas, en déneddylant les CRLs, permettre l'association avec CAND1. Bien que de très notables avancées aient été réalisées ces dernières années dans la compréhension du rôle physiologique exacte du CSN, il reste néanmoins encore beaucoup de travail pour unifier ces données et produire un modèle de leur mécanisme fonctionnel.

Le CSN est un complexe très bien conservé chez les eucaryotes, composé de huit sous-unités (CSN1 à 8 ; de la plus grande à la plus petite des sous-unités) d'un poids moléculaire total de 320 kDa. L'importance de ce complexe CSN dans les cellules eucaryotes se manifeste aussi par les résultats d'études génétiques où le Knock-Out (K.O.) des sous-unités 2, 3, 5 ou 8 chez la souris n'est pas viable au stade embryonnaire. Parmi les sous-unités du CSN, la sous-unité 5, CSN5 (connue aussi sous le nom de Jab1 pour Jun binding activation protein 1) porte l'activité de déneddylation, la fonction la plus étudiée du complexe CSN. Par son rôle central dans l'homéostasie cellulaire, le complexe CSN est souvent associé à la régulation de fonctions cellulaires diverses, comme le contrôle du cycle cellulaire, l'expression génique, l'immunité et la réparation des dommages à l'ADN.

La fonction la plus étudiée jusqu'à présent du complexe CSN à 8 sous-unités correspond à l'hydrolyse de la liaison isopeptidique entre la Culline et Nedd8, qui a, pour conséquence, comme détaillé plus haut, de permettre le turn-over des complexes CRL. CSN5, enzyme de la famille des isopeptidases à zinc, qui correspond au centre catalytique du CSN possède un domaine catalytique dans sa partie N-terminale, le domaine MPN (Mpr1 and Pad1p N-terminal), agrémenté du motif JAMM (Jab1/MPN domain metalloenzyme) qui correspond au motif catalytique. De manière intéressante, l'activité déneddylase portée par CSN5 est dépendante de son intégration au sein du complexe CSN. Prise en dehors de ce complexe, la sous-unité CSN5 est dépourvue de cette activité enzymatique.

De manière notable, le CSN partage avec le chapeau du 26S proteasome et, de manière moins marquée, avec le complexe eIF3, une composition et une architecture communes. Ces trois complexes multiprotéiques contiennent tous six sous-unités comprenant un domaine PCI (Proteasome, COP9, eIF3) et deux sous-unités comprenant un domaine MPN. Leurs activités catalytiques sont liées au système ubiquitine-protéasome. En effet, le chapeau du 26S

proteasome et le complexe eIF3 expriment tous les deux une activité déubiquitylase (DUB), tandis que le CSN a une activité déneddylase. La caractérisation structurale de ces complexes confirme une architecture similaire avec les sous-unités PCI formant de longues ‘pattes’ sur lesquelles sont posées les sous-unités MPN.

Le complexe CSN, mis en évidence il y a 20 ans d’abord chez les plantes, est étudié par de nombreuses équipes dans le monde. Ces efforts globaux se sont principalement axés sur la génétique et la biologie cellulaire du complexe, mais ce n’est que relativement récemment que les aspects moléculaires et structuraux de ce complexe ont commencé à être explorés, probablement du fait de sa complexité. Le projet concerne la caractérisation structurale et moléculaire du complexe, en prenant comme fil directeur la sous-unité CSN5. Cette sous-unité présente un intérêt prononcé du fait de son activité catalytique et du fait qu’elle semble largement impliquée dans différents types de cancers.

## 2. Résultats

### *2.1 L’état inactif de CSN5 dans sa forme isolée est maintenu par l’intermédiaire d’un pont salin.*

La structure cristalline du domaine catalytique de CSN5 a été déterminée en amont du démarrage de ce projet. L’analyse de cette structure cristalline a permis de comprendre pourquoi et comment CSN5 dans sa forme isolée est dénuée d’activité.

L’analyse de la structure de CSN5 confirme qu’elle adopte un repliement de type MPN, complété par des extensions N- et C-terminales qui rendent la surface de CSN5 spécifique. La structure de CSN5 dans sa forme isolée a permis de comprendre les bases moléculaires de sa régulation. Le site actif de CSN5, par comparaison avec d’autres métallo-dépendant isopeptidases et en particulier, la DUB homologue, AMSH-LP (seule enzyme à domaine MPN dont la structure était connue à cette période), est compatible avec la catalyse : la position des résidus du motif catalytique JAMM est compatible avec l’hydrolyse d’une liaison peptidique, comme l’a montrée la comparaison avec l’enzyme AMSH-LP.

Cette analyse ayant montré l’intégrité du site de fixation du zinc, cela laisse donc envisager que l’absence d’activité de CSN5 pourrait être due à son incapacité à recruter le substrat Nedd8. La fixation de celui-ci devra être accompagnée par de vastes changements conformationnels nécessaires pour la création de ce site. Le segment Ins-1 qui correspond à une région impliquée dans le recrutement du substrat dans AMSH-LP est replié sur le site catalytique et fait obstruction à la fixation du substrat Nedd8. Ces observations impliquent que le segment Ins-1 doit changer de position et de conformation pour aboutir à une forme active de CSN5. Allant plus loin dans l’analyse des propriétés de CSN5, nous avons étudié sa flexibilité par dynamique moléculaire. Ce remodelage structural est compatible avec les propriétés de flexibilité et de plasticité de la zone concernée que nous avons exploré *in silico*, montrant une propension de la région Ins-1 à s’ouvrir sur le solvant. Ce travail a permis d’appréhender le degré de flexibilité de cette région de CSN5 et d’attribuer un aspect fonctionnel à cette plasticité. De plus, le fait que le segment Ins-1 adopte des conformations variées dans les différentes structures de protéines à MPN renforce le caractère plastique de cette région de CSN5.

Ce segment Ins-1 de CSN5 présente, au cours des simulations de dynamiques moléculaires, des propriétés de flexibilité élevées de part et d'autre du résidu arginine 106 (R106), qui semble ancrer le segment au site de fixation du zinc via un pont salin avec le résidu acide aspartique 151 (un des résidus en interaction avec le zinc ; D151). La substitution *in silico* de ce résidu confirme à la fois son rôle d'ancre et l'ouverture du segment Ins-1 sur le solvant lors de la perte du pont salin R106-D151. La confirmation expérimentale du rôle de ce résidu dans le mécanisme d'activation de CSN5 a été établie par le fait que le variant R106T de CSN5 est pourvu d'activité isopeptidase. En effet, en utilisant un substrat synthétique soit LRGG-AMC (peptide correspondant aux quatre derniers résidus de l'ubiquitine et de Nedd8 couplé par une liaison isopeptidique à un fluorophore dérivé de la coumarine), soit Nedd8-AMC (substrat synthétique formé de la conjugaison de la protéine Nedd8 avec une molécule fluorescente dérivée de la coumarine), la mutation de l'arginine 106 en thréonine (R106T) confère à CSN5 une nette augmentation de son activité isopeptidasique. La comparaison avec l'activité du complexe CSN entier montre que, bien que il y ait effectivement une activation de l'enzyme par cette substitution, cette activité est relativement modeste par rapport à celle du complexe CSN. Ces observations peuvent être interprétées par le fait que la rupture du pont salin joue probablement un rôle important dans l'activation, mais que le segment Ins-1 peut ne pas se restructurer de manière optimale pour former l'autre moitié du site de fixation du substrat. De plus, il a été montré dans le cadre de nombreuses études que les autres sous-unités du complexe CSN sont elles aussi importantes pour l'expression de cette activité isopeptidase.

## **2.2 L'association de CSN5 en homo-dimère.**

CSN5 forme des dimères de dimères dans le cristal. La présence de dimères en solution a été explorée et vérifiée par différentes techniques de biochimie (gel filtration analytique ; pontage chimique) et de biophysique qui incluent des expériences d'ultracentrifugation analytique (en collaboration avec le Dr. C. Ebel, IBS, Grenoble). L'organisation de CSN5 en homo-dimère en solution pourrait suggérer un nouveau niveau de régulation jusqu'alors ignoré. Du fait de l'intérêt potentiel de l'état d'oligomérisation dans les fonctions et la régulation, nous avons exploré cet aspect dans des extraits cellulaires eucaryotes en collaboration avec le Dr. F.-X. Claret (University of Texas, Etats-Unis). Effectivement, par immuno-précipitation utilisant une double étiquette Flag/His, la présence d'homo-oligomères de CSN5 a été identifiée dans ces extraits cellulaires, laissant envisager la possibilité de la présence de tels assemblages *in vivo*.

L'analyse détaillée de l'interface entre les monomères, grâce à la structure cristalline et par l'intermédiaire de calculs de dynamiques moléculaires, nous a permis de proposer la prépondérance d'un des deux dimères par rapport à l'autre et d'identifier des positions susceptibles d'affaiblir le dimère proposé (contenu dans l'unité asymétrique). Des expériences d'ultracentrifugation analytique et d'immuno-précipitation sur lysats cellulaires confirment, grâce à l'utilisation de variants d'interface, que l'assemblage dimérique prépondérant correspond au dimère A-B observé *in crystallo*. Nous avons mis en évidence la prédisposition de CSN5 à former des homo-oligomères, mais beaucoup reste à faire. En effet, la fonction, le

rôle, la distribution et la prépondérance *in vivo* des homodimères sont encore entièrement inconnus.

### **2.3 Le domaine MPN de la sous-unité CSN6 est capable d'activer CSN5.**

Après avoir examiné les déterminants expliquant le manque d'activité de CSN5 dans sa forme isolée, nous avons exploré l'activation de CSN5 dans le cadre du CSN. Pour cela, nous avons choisi de tester la sous-unité CSN6. En effet, plusieurs études ont montré que CSN5 et CSN6 sont voisins dans le CSN. De plus, une association similaire a été démontrée pour les sous-unités paralagues du chapeau du protéasome, Rpn11 et Rpn8. Après optimisation de l'expression recombinante du domaine MPN de CSN6, nous avons pu montrer que les domaines MPN de CSN5 et de CSN6 sont capables de s'associer et de former un complexe stable sur colonne de gel filtration et par pontage chimique. L'exploration biophysique de ce complexe a permis de montrer que la constante de dissociation de celui-ci est de l'ordre de 6  $\mu$ M pour la forme sauvage, aussi bien que pour le variant R106T constitutivement activée. CSN6 est une sous-unité du CSN qui possède un domaine MPN sans motif JAMM et qui est donc intrinsèquement dénué d'activité catalytique. Ayant montré que les deux domaines MPN s'associent directement, l'effet de CSN6 sur l'état d'activation de CSN6 a été mesuré sur une panoplie de substrats (quatre au total : LRGG-AMC ; Nedd8-AMC ; pro-Nedd8 (précurseur de Nedd8 qui possède une extension C-terminale de cinq acides aminés) ; Nedd8-Cul1/Rbx1 (un des substrats physiologiques de CSN obtenu par neddylation *in vitro* de la sous-unité Cul1 associée à la sous-unité Rbx1)) de nature et complexité variées. Pour l'ensemble de ces substrats, pour lesquels nous avons mis en place et optimisés les conditions de mesure, la sous-unité CSN6 stimule fortement l'activité enzymatique de CSN5. La comparaison du niveau d'activité catalytique atteint par CSN5 en présence de CSN6 avec l'activité du CSN pour trois substrats différents montre que le complexe CSN5/CSN6 possède une activité comparable, si ce n'est légèrement plus efficace que celle du CSN pour le substrat Nedd8-AMC. Bien que cela n'ait pas été anticipé à travers les études décrites dans la littérature, le complexe CSN5/CSN6 possède aussi une activité hydrolytique sur le précurseur de Nedd8, pro-Nedd8, qui approche celle de NEDP1 (NEDD8-specific protease 1 ; aussi connue sous le nom de DEN1 ou SENP8), enzyme à laquelle est généralement attribuée la maturation de pro-Nedd8 dans les cellules. En ce qui concerne le substrat physiologique Nedd8-Cul1/Rbx1, l'hétérodimère CSN5/CSN6 montre une capacité à hydrolyser ce substrat, mais considérablement moindre que celle affichée par le CSN. A travers l'utilisation de ces substrats, une échelle moléculaire a pu être créée et employée pour identifier et caractériser l'activité du complexe CSN5/CSN6. L'ensemble de ces données a conduit à proposer le dimère CSN5/CSN6 comme le cœur catalytique du CSN.

De plus, l'utilisation du variant R106T de CSN5 en présence de CSN6 a mis en évidence un effet synergétique entre la relaxation conformationnelle du segment Ins-1 et l'association de CSN6 concernant l'activation de CSN5.

Il existe d'autres exemples d'association entre une enzyme à domaine MPN et une protéine non-catalytique à domaine MPN dans lesquels la sous-unité catalytique nécessite, pour son activation, la présence de la sous-unité MPN non-catalytique. C'est le cas de la paire



Rpn11/Rpn8 du chapeau du protéasome ou encore de BRCC36 (BRCA1/BRCA2-containing complex subunit 36 ; protéine MPN à activité DUB) qui nécessite soit Abraxas, soit Abro1 dans le cadre des complexes Rap80 ou BRISC, respectivement (BRCC36 isopeptidase complex).

#### ***2.4 L'activation de CSN5 s'accompagne d'une augmentation de l'affinité pour Nedd8.***

Pour mieux comprendre les bases moléculaire et mécanistique de l'activation de CSN5 par CSN6, l'affinité de CSN5 pour Nedd8 a été évaluée par anisotropie de fluorescence, en utilisant Nedd8 étiqueté avec le fluorophore Alexa Fluor 488. L'ensemble des données obtenues montre que (i) le variant R106T de CSN5 possède une affinité augmentée pour Nedd8 par rapport à la forme sauvage, confirmant notre hypothèse que la relaxation conformationnelle du segment Ins-1 est liée à l'interaction avec Nedd8, (ii) l'association de CSN5 avec CSN6 augmente l'affinité pour Nedd8 par rapport à CSN5 seule, (iii) l'affinité pour Nedd8 est comparable pour le variant R106T de CSN5 dans sa forme isolée et pour le complexe CSN6/CSN5, (iv) l'association du variant R106T de CSN5 avec CSN6 induit une augmentation d'affinité pour Nedd8 supérieure à celle du complexe CSN5/CSN6. Pris dans leur ensemble en conjonction avec les données d'activité obtenues, ces résultats suggèrent que le mécanisme par lequel CSN6 contribue à l'activation de CSN5 procède par l'augmentation de l'efficacité de recrutement du substrat, mais que cet aspect n'est probablement qu'une partie du mécanisme d'activation. Il est possible que CSN6 aide à positionner de manière compétente la liaison à hydrolyser par CSN5.

#### ***2.5 Identification de la surface d'interaction du complexe CSN5/CSN6.***

Ayant mis en évidence le rôle central de CSN6 dans l'activation de CSN5, sa structure cristalline fut déterminée par remplacement moléculaire, en utilisant Rpn8 humaine comme modèle. En effet, alors que la structure cristalline d'un fragment de CSN6 de *Drosophila* avait déjà été déterminée, le fait que cette dernière soit particulièrement courte et comprenne de nombreuses boucles désordonnées empêchait son analyse détaillée. C'est pourquoi il a été décidé de produire, de manière recombinante, (la protéine intégrale ne s'exprime pas de manière soluble chez la bactérie) des fragments plus importants de la protéine humaine et de cristalliser l'un d'entre eux. La structure de CSN6 humaine que nous avons déterminée amène de nombreuses informations qui seront largement utilisées dans la suite du projet, comme décrit ci-dessous : (i) cette structure comprend l'intégralité du domaine MPN, contrairement à celle de *Drosophila* et permet ainsi une étude plus fine de la repliement et des propriétés de surface ; (ii) dans la structure de CSN6 humaine, la protéine forme un homo-dimère stabilisé par un motif 'helix swap' où la dernière hélice d'un monomère se fixe dans un sillon de l'autre monomère et *vice versa*. A la fois le mode de dimérisation et la présence de cet échange d'hélices sont tous les deux retrouvés dans l'homo-dimère de Rpn8 humain.

En complément à ce travail biochimique et structural, nous avons entrepris de caractériser le mode d'association de CSN5 et CSN6. Pour cela, des essais de cristallisation extensifs ont été menés, mais n'ont abouti qu'à l'obtention de cristaux de taille et/ou qualité insuffisante pour diffracter les rayons X. En parallèle de ces tentatives, une collaboration avec



la plateforme RMN du Centre de Biochimie Structurale, dirigée par le Dr. A. Padilla, a été établie dans le but de cartographier l'interface d'interaction entre les deux protéines. Très rapidement, l'étude de CSN5 par RMN a dû être abandonnée pour des raisons techniques. En effet, CSN5 a la propension de dimériser, comme développé dans un des paragraphes précédents, ce qui rend difficile l'obtention de spectres RMN exploitables. L'utilisation d'un variant principalement monomérique, bien qu'améliorant la qualité des spectres, n'a pas suffi à régler ce problème. Nous avons donc concentré nos efforts sur CSN6, d'abord par la réalisation de son attribution et ensuite par l'optimisation de conditions expérimentales favorables à l'acquisition de spectres HSQC (Heteronuclear single quantum coherence spectroscopy) pour cartographier la surface d'interaction de CSN6 avec CSN5. L'ajout de CSN5 à CSN6 marqué à l'azote 15 induit de profonds changements dans les spectres HSQC de ce dernier ; ceci inclut le déplacement, l'élargissement, la diminution d'intensité et la disparition de nombreux pics. Le complexe CSN5/CSN6 atteint 50 kDa, ce qui le place à la limite des objets étudiables par la méthode traditionnelle HSQC. Il est donc normal que les spectres HSQC de ce complexe soient difficiles à interpréter. C'est pourquoi et par mesure de prudence que nous avons choisi non pas d'identifier les acides aminés perturbés par l'ajout de la protéine partenaire, comme ce qui est normalement fait, mais, au lieu de cela, d'identifier ceux qui ne subissent pas de modifications dans leurs déplacements chimiques  $^1\text{H}/^{15}\text{N}$ . Cette approche opposée a permis d'identifier deux surfaces potentielles d'interaction. La surface la plus probable d'interaction de CSN6 avec CSN5 fut sélectionnée sur la base de l'analyse de la conservation de séquence de CSN6 et de la surface d'homo-dimérisation utilisée par Rpn8 (paralogue de CSN6 dans le chapeau du protéasome). Nous avons ensuite validé cette interface par mutagenèse et par mesure de l'effet de ces variants sur l'association avec CSN5 et sur l'activation de CSN5. Il est important de noter que cette analyse a mis en évidence que CSN6 interagit avec CSN5 via la même surface que celle utilisée dans le contexte de sa propre homo-dimérisation et que celle utilisée dans le contexte de l'homo-dimérisation de Rpn8. Ceci nous a permis d'en déduire une surface potentielle d'interaction de CSN5 avec CSN6. Ces deux surfaces - celle de CSN6 avec CSN5 et celle de CSN5 avec CSN6 - furent par la suite confirmées par la résolution de la structure du complexe paralogue Rpn11/Rpn8 par deux laboratoires différents.

## ***2.6 Un modèle du complexe CSN5/CSN6 obtenu par amarrage guidé par des données expérimentales.***

Pour pallier la difficulté d'obtention de la structure cristalline du complexe CSN5/CSN6, un modèle moléculaire de cet hétérodimère fut construit par amarrage moléculaire et filtrage des solutions par les données expérimentales.

Pour cela, au vu de l'augmentation de l'affinité de CSN5 pour Nedd8 en présence de CSN6, nous avons décidé d'utiliser un assemblage CSN5-Nedd8 pour les calculs d'amarrage. Effectivement, l'utilisation d'un module CSN5-Nedd8 au lieu de CSN5 seul permet de réduire la surface explorée. Pour la construction de ce module, plusieurs étapes ont été utilisées. La construction du module CSN5-Nedd8 repose sur l'hypothèse que Nedd8 se lie à CSN5 d'une manière similaire à l'ubiquitine distale à AMSH-LP, complexe pour lequel il existe une structure cristalline. Dans un premier temps, nous avons donc vérifié la plausibilité

de cette hypothèse en mesurant l'effet sur l'activité hydrolytique de CSN5 de substitutions à des points de Nedd8 correspondant à des points d'interaction importants entre l'ubiquitine distale et AMSH-LP. L'ensemble de ce travail de mutagenèse a confirmé que les positions importantes de l'ubiquitine distale pour l'interaction avec AMSH-LP correspondent à des positions de Nedd8 qui sont aussi importantes pour la catalyse. Dans une deuxième étape, la conformation du segment Ins-1 que nous avons montré flexible et importante pour l'activation de CSN5 a été modifiée par modélisation pour permettre la fixation de Nedd8 dans le sillon de recrutement du substrat. Les deux partenaires, c'est-à-dire la forme modifiée de CSN5 et Nedd8, ont ensuite été assemblée pour respecter le mode de fixation illustré dans la structure cristalline d'AMSH-LP et ubiquitine.

La construction d'un modèle CSN5-Nedd8 ayant été réalisée, le processus d'amarrage par le programme ZDock a été lancé et analysé. Les 2000 poses ayant les meilleurs scores furent ensuite soumis à une procédure de sélection basée sur les données expérimentales obtenues précédemment. Le tri de ces 2000 solutions a été fortement accéléré par l'utilisation d'un programme informatique d'évaluation spécifiquement conçu et écrit pour cette tâche. Cette procédure a permis l'obtention de modèles compatibles avec l'ensemble des données et très proches de la topologie des homo-dimères des protéines humaines Rpn8 et CSN6.

La publication de deux structures du complexe Rpn11/Rpn8 de levure par deux groupes distincts, concomitante avec notre sélection du modèle le plus plausible du complexe CSN5/CSN6 a permis de comparer le mode d'interaction proposé pour notre dimère avec celui du chapeau du protéasome. Le fragment CSN5 (1-257) utilisé pour la quasi-intégralité de nos études dont les études cristallographiques possède un domaine MPN qui correspond aux résidus 51 à 232 autour duquel s'enroulent les extensions N- et C-terminales. De manière notable, ces extensions sont agrafées par des interactions principalement hydrophobes. Dans les fragments utilisés pour déterminer les structures de Rpn11/Rpn8, les extensions N- et C-terminales sont réduites. La comparaison du mode d'interaction proposé pour CSN5 et CSN6 et observée pour Rpn11/Rpn8 révèle que les deux topologies sont très proches, mais que la présence des extensions N- et C-terminales présentes dans CSN5 empêche les deux topologies d'être identiques. A la suite de cette analyse, une nouvelle procédure d'amarrage fut réalisée en utilisant un fragment de CSN5 débarrassé de ses extrémités N- et C-terminales. L'analyse des solutions les plus probables confirme l'obtention d'un modèle du complexe CSN5/CSN6 dont la topologie générale est identique à celle du complexe Rpn11/Rpn8.

### 3. Conclusion

L'ubiquitylation des protéines et les voies de signalisation qui lui sont liées constituent des processus fondamentaux à l'homéostasie cellulaire. Une des preuves les plus flagrantes de l'importance du système ubiquitine-protéasome réside dans sa dérégulation pathologique récurrente. Le système ubiquitine-protéasome est maintenant le cœur d'efforts globaux pour cibler ses fonctions dans le cadre d'interventions thérapeutiques, comme l'illustre le succès de la molécule bortezomib (ou Velcade) qui inhibe l'activité du protéasome et dont l'utilisation a été autorisée dans le traitement de certaines pathologies hématologiques. Néanmoins la toxicité et l'émergence de résistances contre cette molécule motivent la recherche de dérivés

de celle ci, ainsi que de nouvelles voies d'interventions thérapeutiques. Le ciblage rationnel d'une voie de signalisation nécessite une compréhension détaillée, à différents niveaux de ses fonction et de ses mécanismes de régulation. C'est dans cette perspective de recherche fondamentale pour aider et accompagner de possibles progrès dans le traitement de maladies humaines liées à la dérégulation du système ubiquitine-protéasome que l'étude du CSN s'inscrit.

Au cours de ces trois ans de thèse, le projet de recherche fut principalement centré sur la sous-unité catalytique du CSN, CSN5. Grâce à une stratégie de recherche large incluant des approches de biochimie préparative, analytique, biophysique, bio-informatique, de biologie structurale et de biologie moléculaire, ce projet a permis (i) de caractériser CSN5 dans sa forme isolée, et en complexe avec sa sous-unité régulatrice, CSN6, (ii) de concevoir des outils moléculaires pertinents pour l'étude de ce système et (iii) de mettre en lumière certains mécanismes de régulation et de la fonction de l'activité catalytique de CSN5 et donc du CSN. Ainsi, nous avons établi la raison pour laquelle CSN5 ne présente pas d'activité isopeptidase dans sa forme isolée du complexe CSN, répondant à une interrogation formulée il y a plus de dix ans. Allant plus loin dans la connaissance du modèle, il a été possible de concevoir un variant constitutivement actif de CSN5, ce qui ouvre des perspectives intéressantes dans le domaine de la compréhension cellulaire de la fonction de CSN5 et dans l'optimisation rationnelle d'inhibiteurs de CSN5 avec l'utilisation de ce mutant d'activité. De plus, la sous-unité CSN6 du CSN a été identifiée comme stimulant l'activité catalytique de CSN5. A travers une caractérisation enzymatique détaillée couplée à des mesures biophysiques, il a été possible de proposer un mécanisme d'activation de CSN5 par CSN6, ainsi que l'interface d'interaction impliquée dans cette association. Ce travail de cartographie de l'interface a permis de pallier le manque de données structurales de ce complexe et d'établir un modèle moléculaire obtenu par amarrage guidé par nos données expérimentales.

L'ensemble de ces travaux a permis de se confronter à un domaine biologique compétitif, nécessitant l'utilisation de nombreuses techniques, certaines dans le cadre de collaboration. Mais comme souvent, bien qu'ayant permis de fournir quelques ébauches de réponses à des question biologiques, ces travaux ont aussi abouti à la formulation de nouvelles questions. Et il reste par conséquent de nombreux aspects à couvrir concernant cette protéine de 334 acides aminés. En particulier, notre travail principalement cantonné à une approche *in vitro* devra être prolongé et étoffé par des expériences *in cellulo*. Nous voyons certaines questions biologiques qui devront être examinées et notre travail, ainsi que celui d'autres laboratoires, aura peut-être permis d'aider à commencer à y répondre et/ou à concevoir des outils moléculaires ou chimiques pour y réussir : (i) Quelle est la fonction et l'occurrence cellulaire de la forme isolée de CSN5 ?; (ii) Comment le CSN est assemblé *in vivo* ?; (iii) L'activité catalytique de CSN5 est-elle associée à d'autres fonctions que la régulation des CRLs ?



## Abstract

The COP9 (Constitutive photomorphogenesis 9) signalosome (CSN) is an eight-subunit-containing multiprotein complex (320 kDa) implicated in diverse cellular processes including, cell cycle progression, gene expression and DNA repair via its function in the ubiquitin-proteasome pathway. It is a highly evolutionary conserved protein complex in higher eukaryotes for which its activity is essential. Over years of biochemical and biological studies to elucidate the role of the CSN, its most studied and best understood function is linked to the control of protein ubiquitylation (post-translational modification corresponding to the covalent conjugation of an ubiquitin molecule) by a class of E3 ubiquitin ligases. The CSN exhibits catalytic activity to regulate E3-cullin RING ubiquitin ligases (CRLs) by the removal of an ubiquitin-like protein, Nedd8 (cullin-neural precursor cell expressed developmentally downregulated gene 8), from CRLs. Cycles of neddylation/denedylation are essential for CRL function and the CSN is central in this process through its activity as a CRL deneddylase. Structural and functional similarities link the CSN, the 19S lid of the 26S proteasome and the eukaryotic initiation factor-3 (eIF3). These multi-subunit assemblies comprise six PCI (proteasome COP9 eIF3) domain subunits and two MPN (Mpr1–Pad1–N-terminal) domain-containing subunits. The catalytic activity of the CSN is centred on its subunit 5 (CSN5/Jab1), which hydrolyses the Nedd8-CRL isopeptide bond. CSN5 contains a zinc-dependent isopeptidase catalytic centre constituted of a JAMM (Jab1/MPN/Mov34) motif.

CSN5 incorporation within the CSN complex unleashes its isopeptidase activity, whereas it remains inactive in isolation. The work carried out in the last three years led to several main findings. (i) Having elucidated the crystal structure of CSN5 catalytic domain, biochemical and *in silico* investigations that furthered the understanding of CSN5 molecular regulation led to the identification of a potential molecular trigger enabling CSN5 to be active and the design of a constitutively active CSN5 variant form. (ii) The ability of CSN5 to homodimerise was investigated in solution, *in silico* and in cellular extracts and brought information that could be important for its function. (iii) Further to that work, to address CSN5 activity within the CSN, the contribution of another CSN subunit, mainly CSN6, shown to interact directly with CSN5 was evaluated and this led to the identification of CSN6 as the CSN5 activating subunit. (iv) The biochemical and biophysical characterisation of the CSN5-CSN6 complex was exploited to explore at the molecular level this complex in the context of its binding to Nedd8 and of its integration within the holo-CSN assembly through an integrated approach that includes biochemical, structural, biophysical and computational techniques.

Overall this work allowed to gain an in-depth understanding of the activity determinants and of the regulatory mechanisms that the CSN catalytic subunit CSN5 is subjected to.

**Keywords:** ubiquitin-proteasome system; proteasome; COP9 signalosome; isopeptidase activity; ubiquitin E3 ligase; cullin regulation, MPN, Rpn11, pro-Nedd8

## Chapter 1. INTRODUCTION

### 1.1 Biology:

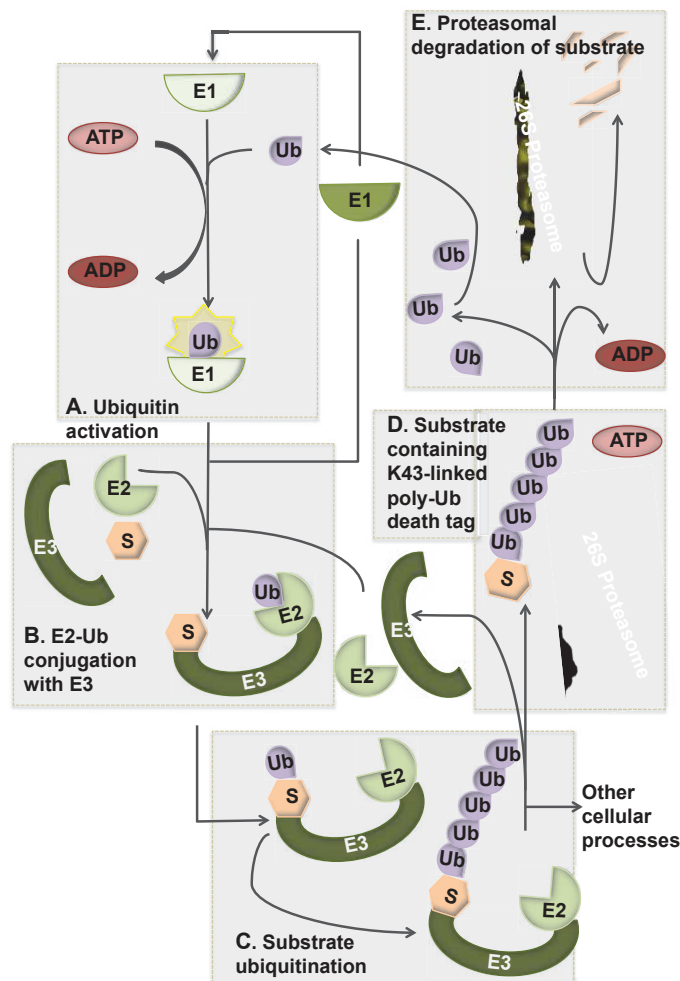
#### 1.1.1 The ubiquitin-proteasome system

Ubiquitin (Ub) is a small and conserved 76-residue (8.5 kDa) protein which is post-translationally conjugated to lysine residues of a wide assortment of proteins. This protein retains a powerful role in mediating and regulating a range of fundamental cellular processes. Although not restricted to this, the best known effect of the Ub molecule is linked to its function as a degradation mark. These include cell cycle progression [1], oncogenesis [2], endocytosis [3] and antigenic peptide production [4]. Ubiquitination, which corresponds to the covalent attachment of an Ub molecule, comprises a three-step cascade and thus implicates three different enzymes (Fig. 1-1). Initially the C-terminal carboxylate of the distal glycine residue of Ub requires activation by sequential adenylation and thioesterification reactions mediated by the Ub activating enzyme, E1. The second step encompasses the transfer of activated Ub from E1 to an activated cysteine residue on the Ub conjugating enzyme, E2, via a transthioesterification reaction. The Ub charged E2 associates with the substrate-bound Ub ligase, E3, and together they collaborate to transfer the thioesterified Ub to the  $\epsilon$ -amino group of lysine residue on the substrate, forming an isopeptide linkage [5-7].

Ub conjugation can either be attached to the substrate as a monomer or form different poly-Ub chains. In the former case, the target protein is mono-ubiquitinated; in the latter case, poly-ubiquitinated. Ub contains seven lysines with the potential to attach to other Ub molecules and the topology of the poly-Ub chains has been shown to generally determine their functional role. For example, modification of a target protein with Ub via Lys48 residue forms an Ub Lys48 chain of generally four or more Ub molecules linked through their Lys-48 residue, which effectively acts as a recognition tag for proteasomal degradation [8]. Whereas, Lys63-linked mono- or poly-Ub chain on the substrate, instead, results in the altered transport or function of the target [9, 10], DNA repair [11], activation of NF- $\kappa$ B [12], and polysome stability [13]. Recent studies have suggested that Ub can also covalently conjugate to N-terminal  $\alpha$ -amino groups via a peptide linkage. This type of Ub conjugation yields to linear poly-Ub chains [14] or N-terminal Ub fusions [15, 16]. The combinatory of ubiquitination modification offers a complex range of regulatory responses.

The Ub-Proteasome System (UPS) is a complicated and highly regulated mechanism discovered in the eighties and for which Hershko, Ciechanover and Rose received the Nobel Prize in 2004. However the complexity of this system has made it challenging to gain

complete understanding. Given the essential role of the Ub system in diverse cellular processes, it is not surprising that its dysfunction contributes to cancer, neurodegenerative and immunological disorders. The first anti-cancerous treatment based on proteasome inhibition that has received approval from the Food and Drug Administration in 2008 under the commercial name of Velcade (Bortezomib), provides a proof-of-principle that targeting the UPS is a therapeutic intervention avenue to further explore [17]. Toxicity effects and the emergence of resistance against bortezomib led to the development of bortezomib derivatives, as well as the identification of other therapeutic intervention lines within the UPS. However novel therapeutic intervention avenues to treat of these diseases would require a better understanding of this system. More specifically, within our work, we focus to gain understanding on the regulation of E3 ligases. Here we discuss key components linked to the controlled mechanism of E3 ligase activation via the neddylation modification.



**Figure 1-1. Schematic overview of the UPS.**

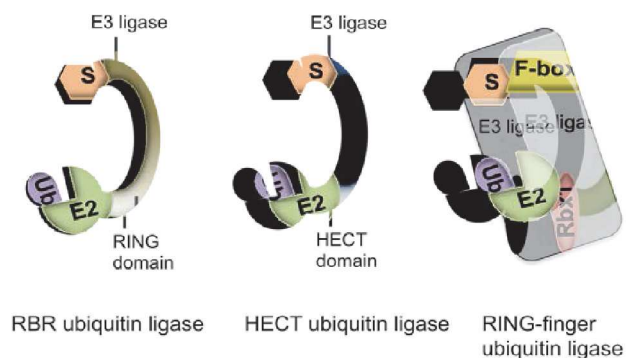
**A.** The three-enzyme cascade for substrate ubiquitination. The first step involves the activation of Ub by E1 (Ub activating enzyme). **B.** The activated Ub conjugates to E2 (Ub-conjugating enzyme) and subsequently transferred to the E3 ligase. The substrate is recruited to the E2-Ub loaded-E3 ligase. **C.** Ub is transferred from the E2 to the substrate. Several cycles of E1-E2-E3 reactions allow the poly-ubiquitination of the substrate. **D.** A poly-ubiquitinated substrate is released from the E3. Depending on the Lys linkage the proteasome recognises the poly-Ub chain as a signal to de-ubiquitinate and destroy the substrate protein. **E.** The unfolding of the substrate by the proteasome is an ATP-dependent process that includes the removal of the Ub chain through the hydrolase activity of the 19S proteasome lid. The substrate is then threaded and unfolded into the proteasome chamber where the proteases are located. Ub molecules and amino acids can be recycled.

### 1.1.2 E3 ligases



Mammalian cells express one or two E1 activating enzymes, 37 E2 conjugating enzymes and more than 600 E3 ligases. The E3 ligase is responsible for binding to both the E2-Ub complex and the target substrate. It functions to catalyse the transfer of Ub from the active site cysteine of the E2 enzyme to the substrate lysine residue or the N-terminus [18]. As such the E3 ligase contains dual roles, both molecular matchmaking and catalysis (or enhancement of Ub transfer to substrate), assembling the appropriate E2 with the correct substrate and vastly increasing the rate of Ub transfer. The increase in rate of Ub transfer by E3s was found to range up to 87-fold in the case of the E2, UbcH5b by the E3, C-NOT4 [19]. The presence of a large pool of E3 ligases accounts to the requirement to specifically target a broad array of substrates, although the determinants and regulation of substrate recognition are not yet entirely known [20].

The E3 ligase family is commonly classed into three different types, namely the RING (really interesting new gene)-finger, RBR (RING-between-RING) and HECT (homologous to E6-associated carboxyl terminus), each of which is characterised according to their conserved structural domains and the mechanism, which they employ to mediate Ub transfer from the E2 to the substrate (Fig. 1-2). The general role of E3s is to bring an E2-Ub into the proximity of the target protein. However different types of E3 ligases use different molecular interactions. The RBR and HECT subfamilies utilise a two-step reaction to ubiquitinate the substrate, in which the Ub is initially transferred from the E2 to an active site cysteine located on the E3 and the Ub is then moved from the E3 to the target protein [21, 22]. The RING subfamily members however bind simultaneously to both the E2-Ub thioester and the substrate and are thus able to directly catalyse the transfer of the Ub from the E2 enzyme to the substrate [18, 23].



**Figure 1-2. E3 ligases.**

Classes of E3 ligases: RBR, HECT, RING-finger. The RBR type, a monomeric E3 contains a RING-finger domain that binds to E2s. The HECT E3s contain a HECT domain that associates with E2. RING-finger E3s are multimeric RING-finger E3s that bind to E2 enzymes via their RING-finger domain, Rbx1. This type of E3s also contains the scaffold protein Cullin and a F-box protein that binds to an adaptor protein and recognises the substrate.

The RING-finger domain is a subtype of zinc-finger domain often found in proteins in mammals [24]. Around 0.7 to 0.75% of the mammalian genome encodes for RING-finger-domain proteins where more than 600 of them are potential RING-finger E3s [20, 25]. Within the E3 ligases, RING-finger proteins constitute the largest class and can further be divided into two major sub-classes: simple RING-finger ligases and Cullin-based RING-finger ligases (CRL). In the simple RING-finger E3s, the RING-finger and the substrate-binding domain are located within the same polypeptide, whilst CRLs are multiprotein complexes and utilise RING-box protein 1 (Rbx1) or Rbx2, in complex with a Cullin scaffolding protein and a



substrate receptor module (that comprises either one polypeptide or two). The CRLs are thus heteromeric enzymes containing the scaffolding subunit, a RING-box protein and a bridging protein known as the F-box protein. The RING-box contains a RING-like domain and is responsible for both binding to the E2 conjugating enzyme and also simulates Ub transfer, whilst the F-box is the site for substrate binding [26, 27].

### 1.1.3 Cullin-RING Ub ligases

The multimeric CRL subclass comprises of up to 240 enzymes that exert a profound effect on eukaryotic cells and organisms through the regulation of proteins key to cell homeostasis. Approximately 20% of the total protein degradation performed in cells by the UPS is accounted to CRLs [28]. Additionally CRLs are also implicated in numerous processes that underlie normal development and physiology through ubiquitination reactions that are not linked to target protein degradation [29]. Much of the architecture and regulation of the CRLs have been contributed by the laboratory of Professor B. Schulman (St. Jude Children's Research Hospital, USA). The Cullin subunit of CRLs, an elongated protein, forms the backbone of this multisubunit complex, on which are recruited the other subunits. The C-terminal end of Cullin is comprised of a globular domain that wraps itself around the RING protein subunit. In turn, the RING protein binds the E2-Ub complex to form the enzymatic core. On the elongated Cullin structure, the long stalk-like N-terminal part consisting of several Cullin repeats utilises adapter proteins to recruit target substrate(s). The most structurally and functionally characterized mammalian CRL is the SCF (Skp1 (for S-phase kinase-associated protein 1)-Cullin1-F-box) Ub ligase. This E3 ligase comprises of Cullin1 (Cul1; there are seven different Cullins, from Cul1 to Cul7 in human), the RING domain protein Rbx1 (also named Roc1 or Hrt1), the adapter protein Skp1 and a substrate receptor protein (F-box protein), where examples include Skp2 (for S-phase kinase-associated protein 2) or  $\beta$ -TrCP (for  $\beta$ -transducin repeat-containing protein). A common F-box motif is shared between the substrate receptor proteins of the SCF complexes mediating their association to Skp1. Substrate specificity of the SCF is largely mediated through the F-box protein, which commonly recognises phosphorylated target proteins [30], whilst the Cul1-Rbx1 harbours the core ligase activity by catalysing the Ub transfer from the E2 enzyme to the substrate [31].

Given their predominant roles in controlling myriad of cellular processes and the profound dysregulation of some elements of CRLs in proliferative diseases [32], there is a strong rationale in gaining insight in the function and regulation of CRLs. To date, all CRLs appear to be under the control of a number of common regulatory mechanisms [24]. The most documented regulatory mechanism of CRLs accounts for the conjugation of the Cullin subunit with an Ub-like protein, Nedd8 (for Neural precursor cell expressed, developmentally down-regulated 8; known as Rub1 in plants and budding yeast) [33]. It is now well established through numerous biochemical and structural studies that CRLs are covalently modified with a Nedd8 protein at a conserved lysine residue located at the C-terminus of the Cullin subunit. Neddylation of Cullin proteins is carried out via an analogous pathway to the ubiquitination of proteins. A Nedd8-specific activating enzyme, known as NAE, activates

the Nedd8 protein which is then transferred to a Nedd8-conjugating enzyme, Ubc12, which associates with substrate specific Nedd8-E3 ligase to covalently conjugate the Nedd8 to the target substrate [34, 35]. Rbx1 (Roc1), the RING domain-containing protein accountable for the E3 Ub ligase activity is also necessitated for the neddylation of Cullins [36]. *In vivo*, the neddylation of CRLs is required for their E3 ligase activity and is suggested to play a role in E2-Ub complex recruitment to the E3 ligase complex and thus stimulating the formation of poly-Ub chains [37-39]. It is also thought that CRL neddylation could promote the selective recruitment of specific substrates, such as Ubx7 [40, 41]

Another regulatory mechanism of Cullin-based E3 ligase activity implicates the CAND1 (for Cullin-associated NEDD8-dissociated protein 1) protein. This protein binds to the Cullin subunit and specifically associates with the deneddylated form of the protein [36, 42-44]. Thereby the binding of CAND1 to Cullin disrupts the CRL assembly and, in particular, the association of the Cullin subunit with the adaptor and substrate recognition subunits. This consequently results in the inhibition of CRL E3 ligase activity. Although detailed insights into the interaction between Cull1 and CAND1 has been reported [44], the mechanism by which CAND1 competes with the substrate recognition subunit is not yet fully understood [44-46].

Moreover, the COP9 (for COnstitutive Photomorphogenesis 9) signalosome (CSN), a highly conserved multiprotein complex, also interacts with CRLs and functions to deneddylate neddylated CRLs, cleaving the Nedd8-Cullin isopeptide bond [47, 48]. The requirement of Cullin neddylation of E3 ligase activity and the implication of the CSN complex suggested that, *in vivo*, Cullin proteins undergo complex cycles of neddylation and deneddylation [24]. The CSN complex and its contribution in the UPS are further detailed in the sections to follow.

#### **1.1.4 Proteasomal degradation**

The 26S proteasome is a highly conserved 2.5 MDa-protein complex comprising two copies each of at least 32 different subunits. This protease machinery functions primarily to degrade proteins that have been modified by the attachment of poly-Ub chains. The overall structure of this multi-protein complex can be divided into two major subcomplexes: the 20S core particle containing the protease subunits and the 19S regulatory particle, which regulates the function and activity of the CP. Structurally the core particle forms a barrel-shaped assembly composed of four rings, each of seven subunits. The proteolytic active sites, facing into a sequestered proteolytic chamber, are located on the two inner  $\beta$ -rings [49, 50]. The isolated internal chamber ensures that protein substrates can access it only after their deubiquitination, unfolding and translocation performed by the regulatory particle. The outer  $\alpha$ -rings of the 20S core particle accommodate either one or two regulatory particles to form the holo-26S proteasome. The 19S regulatory particle is further composed of two different multisubunit substructures, known as the base and the lid [51]. The base of the 19S regulatory particle

contains 6 homologous ATPases, Rpt1 through to Rpt6 and 4 non-ATPases subunits, Rpn1, Rpn2, Rpn10, and Rpn13. The ATPases form a direct interaction with the  $\alpha$ -subunit of the core particle and function to unfold substrates and are required for channel opening before the translocation of substrate through the gated channel into the 20S cavity. Rpn10 and other Ub receptor proteins, including Rad32 and Dsk2, that reversibly interact with the proteasome are implicated in delivery the ubiquitinated substrate to the proteasome. The second substructure of the regulatory particle, the lid, is a 400 kDa complex, composed of eight subunits, Rpn3, Rpn5 to Rpn9, Rpn11 and Rpn12. Six of the lid subunits contain a PCI (for Proteasome, COP9, eIF3) domain, whilst the other two contain a MPN (for Mpr1 and Pad1p N-terminal) domain. The lid is responsible for cleaving the Ub chain from the substrate via its deubiquitinating activity.

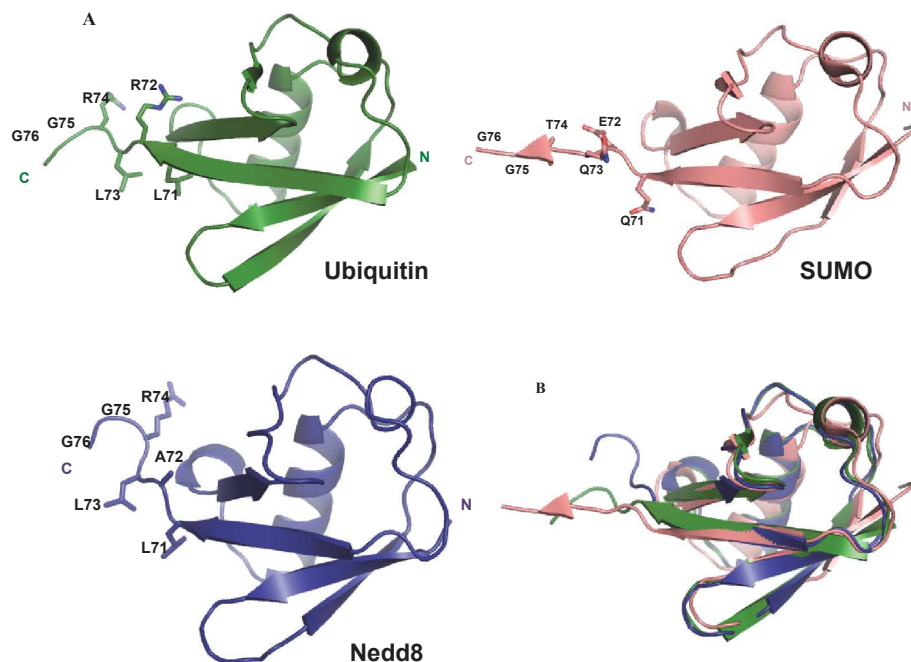
In eukaryotes, the proteasome-dependent degradation pathway is essential for the elimination of damaged or misfolded proteins, the degradation of short-lived regulatory proteins and the definition of signalling rhythms, such as the cell division cycle, during which specific proteins are degraded at specific times. The proteasome operates via a multistep mechanism, with peptide bond hydrolysis being only the last step in this complex machinery. Additionally critical upstream processes include substrate, recognition, deubiquitination, unfolding and translocation. These processes take place at a distance from the proteolytic core site and are mediated by the regulatory particle subcomplex. Proteasomal substrates are initially selected by the 19S regulatory particle via the recognition of the poly-Ub chains attached to the substrate protein destined to be degraded. Several subunits or proteins associated to the lid specifically recognise Ub, but a complete understanding of the role of each of these is yet to be unravelled. However, it is clear that the Rpn11 subunit of the proteasome lid plays an essential role in protein degradation. Indeed, following the selection of the labelled substrate, the poly-Ub chain is cleaved off by the deubiquitination activity of the lid subunit, Rpn11. The substrate protein is then unfolded and translocated into the 20S cavity where it is hydrolysed.

#### **1.1.5 Nedd8 - An Ub-like protein**

The timely, transient and subtle regulation of cellular processes including, for example, protein activity, function, stability, and localisation is required for cellular homeostasis. Post-translational modifications, reversible in nature, are a widely used strategy for these controls. In some cases, post-translational modification may include a small chemical substitution, as is the case for phosphorylation, acetylation, and methylation, whilst, in others, it includes the covalent attachment of another polypeptide chain. Ubiquitination, for example, also described in earlier sections, corresponds to the latter, a post-translational modification with the covalent attachment of an Ub molecule to a target protein. Ub is the first described molecule of this type, but since its discovery, many other proteins that share its distinctive structure, known under the collective name of Ub-like molecules (Ubl) have been reported. The covalent modification of target proteins by these Ubl is multiple. For example sumoylation and neddylation refer to the covalent modification of proteins by one of the SUMO (Small

Ubiquitin-related MOdifier) proteins and by Nedd8, respectively. Post-translational modification of proteins by Ub and Ubl is an essential mechanism regulating various cellular processes (Fig. 1-3A).

In eukaryotes, the conserved *NEDD8* gene is encoded as a non-conjugatable precursor, named pro-Nedd8, that contains one or more additional residues (depending on the considered organism; 5 amino-acid extension in *Hs* and in *Mus musculus* (*Mm*)) in C-terminal of its functionally mature C-terminus at position 76. In their immature form, pro-Nedd8 requires proteolytic processing of its C-terminus to generate mature Nedd8 exposing a C-terminal glycine residue essential for the conjugation of Nedd8 to its substrate protein. The expression of a non-functional precursor is a feature also found in Ub, as well as in other UBLs, where their C-terminal extension is proposed to serve as a prevention of unprocessed proteins from entering the conjugation pathway. However, experimental evidence supporting this hypothesis is still lacking. Nevertheless, precursor processing is strictly required for the recognition of Nedd8 by the E1 enzyme responsible for Nedd8 activation, making Nedd8 processing enzymes essential for Nedd8 function. The Nedd8 processing activity is carried out by members of the Ub C-terminal hydrolases (UCH) subfamily from the family of deubiquitinating enzymes (DUBs). Members of the UCH subfamily, Yuh1 in yeast and UCHL3 in human and mouse, possess dual specificity by processing both Nedd8 and Ub [52]. NEDP1 (Nedd8-specific protease 1), also known as DEN1 and SENP8, is a member of the C48 peptidase family of proteins and is the sole isopeptidase known to date to function exclusively in Nedd8 processing and deneddylation [53].



**Figure 1-3. Structure of Ub and Ubl proteins.**

**A.** Crystal structures and Ub (PDB code: 1UBQ), Nedd8 (PDB code: 1NDD) and SUMO (PDB code: 1TGZ), all containing a common secondary structure. The last 5 C-terminal residues are presented in

stick mode. **B.** Superimposition of Ub and Ubl proteins. High spatial homology is observed for all proteins.

### 1.1.6 The CSN

The CSN was initially discovered upon mutant screening in *At* for candidates displaying a constitutive photomorphogenic phenotype [54]. This described a phenotype where plants grown in the dark resembled those cultivated under light conditions. Following this observation, cloning of all the genes implicated in this process was performed and indicated that they are extremely conserved from fission yeast to humans [54]. Most interestingly, the CSN complex is an eight subunit complex, CSN1 through CSN8, (named according to the order of descending molecular weight), where all eight subunits show high homology to the eight subunits of the 19S proteasome lid and more distantly to the eukaryotic translation initiation factor 3 (eIF3) [54].

Structurally two distinct domains can be defined in the CSN and similarly in the 19S proteasome lid and the eIF3 complex. These include the PCI and the MPN domains [48]. Six CSN subunits (CSN1-4 and CSN7-8) contain a PCI domain, which has been characterised to hold N-terminal bihelical repeats followed by a winged-helix (WH) subdomain [55-58]. Previously suggested by cryo-electron microscopy (EM) data and more recently shown by the crystal structure of the human CSN complex, the PCI proteins arrange to form an open ring structure by associating through their WH subdomains [46, 58-61]. In the CSN complex, the N-terminal helical repeats of some of the PCI subunits function in accommodating the SCF substrate. In particular, subunits CSN1, CSN2, CSN3, CSN4 and CSN8 are thought to be the main contributors in the interaction of the CSN with SCF [46, 61]. Additionally, PCI proteins are considered to be important for maintaining the integrity of the complex via protein-protein interactions. Further roles of these PCI-containing subunits have been proposed, particularly via the recruitment of associated activities, such as protein kinases (examples include, protein kinase D, AKT, casein kinase 2) [62]. The physiological role and regulation of these associated activities are yet to be documented, but it appears that the biology of the CSN is not limited to its eight subunits.

The remaining two subunits of the complex, CSN5 and CSN6, are MPN domain-containing proteins exposing a metalloprotease fold [63-66]. The two MPN subunits share sequence and structural similarities, however only CSN5 contains the complete zinc-metalloprotease motif, known as the JAMM (Jab1 (for JUN activating binding protein 1)/MPN domain metalloenzyme) motif. As such findings have implied that two kinds of MPN domains are present; a catalytic one that coordinates a  $\text{Zn}^{2+}$  ion, known to have a MPN<sup>+</sup> domain, and another, void of catalytic activity referred to as a MPN<sup>-</sup> domain [67]. The JAMM motif of CSN5 is essential for pursuing isopeptidase activity and constitutes the catalytic centre of the CSN complex. MPN domain-containing subunits comprise a N-terminal MPN fold that precedes a helical bundle. A similar organisation has been described for the 26S proteasome lid and the eIF3 complex.

Initial insight on the molecular basis of the CSN complex function emerged with the identification that, in human cells, the CSN forms a stable complex with CRL. This stable association was shown in the context of the SCF complex [47]. Similarly, in *At*, the same interaction between the two multiprotein complexes was detected [68]. To date, the most documented function of the CSN complex is attributed to the deneddylation of Cullins, ie the removal of the Nedd8 modification of the Cullin [47]. However, this function might not be limited to the actual deneddylation of Cullins, but in the light of recent discovery [46] could extend to the non-enzymatic sequestering of CRLs.

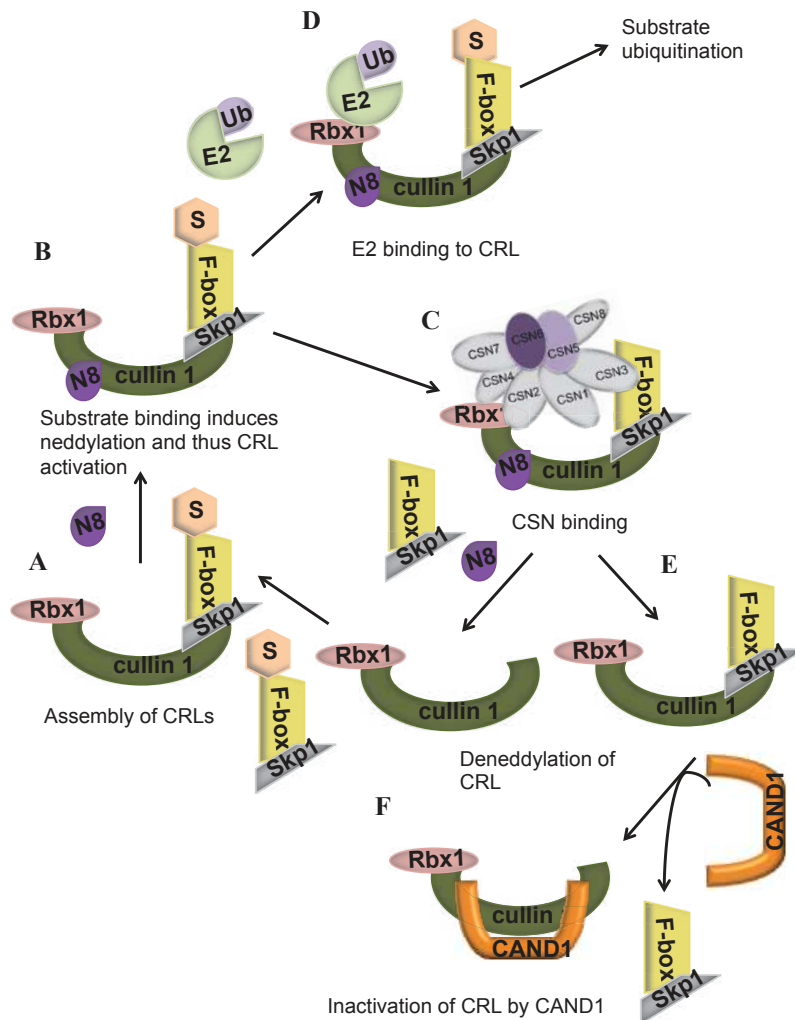
#### 1.1.7 SCF regulation by the CSN complex

The CSN complex forms a stable complex with SCF and accounts to its regulatory mechanism. The CSN was shown to play a critical role in controlling the modification of the Cul1 subunit of the SCF [47]. Whilst approximately 10-20% of Cul1 are typically modified by Nedd8, CSN-deficiency in cells results in the complete neddylation of the entire population of Cul1 and in reduced SCF activity [69]. The reason for this outcome was identified to be that CSN-deficient cells lack the enzymatic activity required to cleave the Nedd8 from the Cul1, thus to deneddylate. It was subsequently shown that this enzymatic activity resides on the 5<sup>th</sup> subunit, CSN5, of the CSN complex [70]. Today it is widely considered that the CSN complex primary function is to mediate deneddylation of Cullins and appears to be the most efficient enzyme for this job, if not the only one [33, 71]. The interplay between the CSN and CRLs determines the activity and specificity of CRL ubiquitination and contributes to the correct functional cycle of CRLs (Fig. 1-4). This mechanism catalyses the specific ubiquitination of substrates that are after subsequently degraded by the 26S proteasome. *In vitro* the process of ubiquitination is stimulated upon the neddylation of the Cullin subunit, whilst the removal of Nedd8 carried out by the CSN5 subunit of the CSN complex, deactivates the CRLs. Beside the inhibitory effect of the CSN within this catalytic mechanism observed *in vitro*, further CSN-mediated inhibition of CRL function is also observed in a noncatalytic mode [46, 72]. This mode of inactivation carried by the CSN prevents the CRL substrate adaptors from auto-catalytic degradation and enables the reassembly of the CRLs, maintaining their optimal activity [45]. Cullin deneddylation is thought to be an essential step for the CRLs to display the various substrate specificity necessary for controlled protein substrate ubiquitination. The stability of the complex formed between the CSN and the CRL is further dependent on substrate availability [46, 73, 74]. This has been demonstrated in the case of the natural cell cycle inhibitor, p27<sup>Kip1</sup> (also known as CDKN1B (for Cyclin-dependent kinase inhibitor 1B); referred to as p27 in the rest of the manuscript), which, upon its phosphorylation stimulates the activity of SCF thus controlling its own ubiquitination and degradation [74].

Various mechanisms have been proposed to explain the positive regulation of CRL activity by the CSN that has been unravelled *in vivo* and that is in sharp contrast with *in vitro* results. As



mentioned above, studies have shown that through Cullin deneddylation and subsequent CRL inactivation, the CSN prevents the auto-ubiquitination of a number of Cullin substrate-receptor subunits in the absence of bound substrate [33, 45]. The requirement of both Nedd8 conjugation and deconjugation for optimal CRL function suggests that *in vivo* CRLs undergo rapid cycles of neddylation and deneddylation [75]. Based on this hypothesis, a model of the CRL functional cycle has been developed, in which substrate binding to the CRL complex induces Cullin neddylation, thereby resulting in CRL activation. Following substrate ubiquitination and dissociation, the deneddylation event performed by the CSN may complete this functional cycle that will resume upon the binding of a new substrate. Cullin deneddylation would protect CRL against auto-catalytic ubiquitination and promote the dissociation of the CRL substrate-receptor subunit. Moreover, it has also been suggested that the CSN could be implicated in the regulation of CAND1 binding [33, 75]. CAND1 interacts exclusively with Cullin subunits in a deneddylated form. Additionally, the binding of CAND1 to the Cullin subunit and the binding of substrate-receptor subunit are mutually exclusive. As such, the deneddylation activity of the CSN may function to promote CAND1 binding to Cullin and further prevent CRL disassembly mediated by CAND1 [45]. Furthermore, the CSN may play a regulatory role in the dissociation of poly-ubiquitinated substrate proteins from the CRL complex [76]. Lastly, it was proposed that the CSN might intervene in recruiting critical binding proteins to CRL complexes. It has been repeatedly reported that the CSN associates with various protein kinases that may arbitrate essential functions in the CRL activation cycle including the exchange of substrate-receptor subunits. As such, deneddylation by the CSN may be crucial for promoting its own dissociation to enable efficient progression through CRL functional cycle [77].



**Figure 1-4. CRL functional cycle.**

**A.** Substrate binding to the CRL complex induces Cullin neddylation. **B.** Nedd8 binds to and activates the CRL complex. **C.** and **D.** Either the neddylated CRL complex binds to the CSN complex (C) or the E2-Ub complex associates with the RING-finger protein, Rbx1, and ubiquitin transfer to the substrate is performed (D). **E.** The CSN deneddylates the CRL complex resulting in loss of CRL activity. The CRL complex is disassembled, substrate binding is required for the neddylation cycle to restart. **F.** CAND1 binds to the Cullin by competing off the F-box and Skp1 proteins.



## 1.2 Cancer:

### 1.2.1 The CSN is overexpressed in cancers

Deregulation of the CSN complex or of its interactions with other proteins can affect one of many cellular functions, including DNA repair, cell cycle control, angiogenesis, and microenvironmental homeostasis, all of which appear crucial for tumour development [78, 79]. In the recent years, accumulating evidence suggests a correlation between the aberrant functioning of the CSN complex and its significant role in multiple cancers. The CSN and its implication in cancers have readily made it an attractive target for therapeutic intervention [79, 80]. In many types of cancers, including glioblastoma, breast cancer, myeloma, and leukemia, frequent overexpression of CSN subunits has been identified. Overexpression has been most evidently observed for the MPN domain-containing proteins, CSN5 and CSN6, of the complex, which, in a variety of cancers, appears correlated with cancer progression and poor survival [81-84], whilst inhibition or knockdown of the CSN5 subunit resulted in suppressed tumour growth in mice [81, 85] and in human cancer cell lines [85]. In-depth understanding of the processes that regulate the mode of action of the CSN complex is required for targeting this complex for cancer intervention. Considering the essential cellular functions associated to the CSN, it is reasonable to assume that potentially multiple regulatory mechanisms exist that allow its coordinated action. The understanding of such regulatory mechanisms at the molecular detail remains unclear, however recent structural findings will provide valuable insight into the regulation of this essential complex in the protein degradation pathway.

Given its critical link with the Ub-mediated protein degradation pathway, the CSN complex is implicated in the degradation of various tumour suppressor and oncogene products [79]. These include p27 [86, 87], p53 [84, 88], Mdm2 (for Double minute 2 protein) [84, 89], Smad7 (for Mothers against decapentaplegic homolog 7) [90], Runx3 (for Runt-related transcription factor 3) [91], Id1 (for Inhibitor of DNA binding 1) [92], Skp2 [93] and HIF-1 $\alpha$  (for hypoxia-inducible factor 1- $\alpha$ ) [79, 94].

CSN6 has been an interesting candidate due to its involvement in signalling pathways during tumorigenesis. Although the detailed mechanism remains unclear, CSN6 is thought to be implicated in carcinogenesis and tumour development. Overexpression of this subunit is revealed in many types of cancer including, glioblastoma, breast cancer, myeloma and leukemia, where amplification of the CSN6 genomic region has been detected [80]. Studies have also shown that the gene amplification of CSN6 in breast cancer is positively correlated with tumour size. Human CSN6 was also identified to bind with the human immunodeficiency virus 1 accessory protein Vpr (hVIP) and found to be implicated in cell proliferation and the G2/M phase transition of the cell cycle [95]. An additional link has

been identified between the RING domain-containing p53 ubiquitin ligase, MDM2 and CSN6. CSN6 was also shown to be implicated in the COP1 driven ubiquitination and proteasomal degradation of p53.

Although overexpression of CSN5 and CSN6 has been documented and is detailed here in the following sections, little is known about the implication of other subunits in cancer. For example, the overexpression of the CSN2 subunit can lead to VEGF production [96] and, by affecting the anaphase-promoting complex, can increase the stability of cyclin A, the activatory subunit of cell cycle cyclin-dependent kinases (CDKs). CSN2 overexpression can further result in chromosome instability [97], however the *CSN2* gene is lacking in numerous cancers [98]. The CSN3 subunit is found to be involved in myeloid leukemia factor 1-mediated growth arrest [99]. Deficiency of CSN3 leads to dampened p53 activation, easing cell proliferation and affecting COP1-mediated p53 degradation [99]. However increased cell death in *Csn3*-knockout mouse embryos where reduced CSN3 expression was observed and correlates with enhanced survival in cancer patients carrying the Ras oncogene [100]. Moreover, *Csn4* deficiency was reported to downregulate Skp2 which is implicated in p27 degradation [93, 101]. Although CSN subunits could have roles in cancer, further characterisation of their implication is required.

### 1.2.2 The p27 axis

Through CSN function as a direct regulator of SCF, several essential products of oncogene or tumour suppressor genes appear to be deregulated as a result of CSN overexpression. One such example includes the function and regulation of the tumour suppressor p27. p27, a client of SCF<sup>Skp2</sup> is a key protagonist in the control of cell cycle, as a natural inhibitor of the proliferative activity of cell cycle CDKs. Its controlled degradation triggered by CDK2-dependent phosphorylation and subsequent ubiquitination is permissive for cell cycle progression. p27 is often downregulated in human cancers and it has also recently been found mutated [102, 103]. It functions to protect healthy cells from undergoing anarchical cell proliferation. Given the central role of p27 in cell proliferation control, many regulatory mechanisms target this protein, among which, phosphorylation and sub-cellular localisation. Deregulation of p27 affects the stability of this protein which is required for the maintenance of its tumour suppressor function and would thus consequently cause an impact on tumorigenesis. Ubiquitination of p27 is carried out by SCF<sup>Skp2</sup> ubiquitin ligase containing the Skp2 substrate-recognition component necessary for associating with p27 [104]. Several lines of evidences link p27 downregulation and the CSN. The human CSN complex together with SCF<sup>Skp2</sup> is implicated in regulating the degradation of p27 through ubiquitination [93]. Interestingly, studies have now shown that p27 downregulation is correlated with CSN5 overexpression in cancers [105, 106]. The CSN5 subunit has also been identified to bind directly to and to regulate the export of p27 from the nucleus to the cytoplasm [107]. Resistance to CSN5-mediated degradation of p27 is gained upon CSN5 knockdown [108] or deletion of the CSN5 binding domain [108]. In which case, modified p27 proteins can then efficiently inhibit HER2-activated cell growth, CDK2 activity, cell proliferation and

transformation [108]. Although the contribution of CSN5 has been characterised, the general mechanism of CSN5-mediated p27 downregulation, as well as the implication of other CSN subunits is unknown.

### 1.2.3 The p53 axis

Recent studies have shown that both MPN domain-containing proteins of the CSN complex, CSN5 and CSN6, are involved in the regulation of p53 levels in cells, by controlling p53 ubiquitination. p53 plays a central role in the regulation of cell cycle progress upon DNA damage detection. Indeed p53 acts as a G1/S checkpoint, halting the cell cycle to allow repair time and thus prevents mutations from being propagated. This function of p53 warranted its name 'The Guardian of the Genome'. *TP53*, corresponding to the gene coding for the p53 protein is the most commonly mutated gene in human cancers [109]. Mutations of *TP53* occur throughout the gene product and results in a loss or alteration of the function/regulation of p53. Factors that control p53 function, such as CSN5 and CSN6 are therefore of upmost importance to understand the complexity of p53 biology. However, at the moment, only little is known of these functional relationships, as described below.

Initial findings indicating the implication of CSN5 in the stability regulation of p53 were obtained from experiments using *Csn5*-null mice embryos highlighting an accumulation of p53 protein levels [110]. Additional evidence suggesting a link between CSN5 and p53 was also found from experiments focused on the CSN-associated kinase inhibitor, curcumin (chemical IUPAC name: (1E,6E)-1,7-bis(4-hydroxy-3-methoxyphenyl)hepta-1,6-diene-3,5-dione), a natural pigment isolated from the plant *Curcuma longa* [89]. Currently, curcumin in a combination therapy with gemcitabine, and celebrex, is used in phase III clinical trials as treatment for pancreatic cancer therapy. Although the tumour suppressive effect of curcumin has been reported, the way this effect is mediated remains unclear. It was suggested that curcumin downregulates both CSN5 and p53 [89], but the molecular mechanism of this effect is currently not known. Further biochemical studies indicate that CSN5 interacts with p53 through its MPN domain which is sufficient for p53 downregulation [89]. Mdm2 is the E3 ligase responsible for the ubiquitination of p53. Studies have shown that *MDM2*-null cells also display lack of CSN5-mediated p53 degradation thus indicating that Mdm2 is required for the degradative function of CSN5 [89]. Moreover, through an unknown mechanism, CSN5 expression increases Mdm2 stability by reducing MDM2 self-ubiquitination [111, 112]. Interestingly, as for other proteins such as p27, CSN5 can mediate the translocation of p53 from the nucleus to the cytoplasm, thereby facilitating p53 degradation [89, 113]. As such it can be concluded that CSN5 is implicated in the enhancement of nuclear export and degradation of p53 via Mdm2-mediated p53 ubiquitination leading to the impairment of DNA damage checkpoint.

Similarly, the comparison of expression levels of Mdm2 and CSN6 in normal and cancerous breast tissues revealed that the expression of CSN6 was concurrently overexpressed with Mdm2 in these tissues [84]. Mechanistically Zhao et al. proposed that CSN6 could prevent

Mdm2 auto-ubiquitination at lysine 364 and thus results in its stabilisation and subsequent degradation of p53. Further, they claim that the lack of p53 degradation observed in *Mdm2*-null mouse embryonic fibroblasts (MEFs) could be indicative that CSN6-mediated degradation of p53 is Mdm2-dependent [84]. These results suggested that CSN6 is an oncogene with positive activity toward Mdm2 function within the MDM2-p53 signalling pathway. CSN6 has also been reported to interact with another p53-specific E3 ligase, COP1. This E3 ligase targets p53 for degradation by the UPS and is therefore reported as a negative regulator of p53 and thus maintains low levels of p53 in unstressed cells and inhibits p53-dependent transcription and apoptosis events [114]. Besides interacting with COP1, CSN6 was also reported to be involved in the degradation of the 14-3-3 $\sigma$  gene product (a gene that is upregulated by p53) [115]. In a manner reminiscent to that proposed for Mdm2, CSN6 was found to stabilize COP1 by reducing COP1 auto-ubiquitination and thus decreasing the rate of COP1 turnover [115]. Two putative physiological effects of CSN6 in the COP1-p53 pathway have been suggested. These include that (i) the stabilisation of COP1 essentially enhances COP1-mediated p53 ubiquitination and subsequent degradation and (ii) the implication of CSN6 results in the degradation of the 14-3-3 $\sigma$  protein.

#### **1.2.4 The p57 axis**

Recently it was reported that the CDK inhibitor, p57<sup>Kip2</sup> (also known as CDKN1C (for Cyclin-dependent kinase inhibitor 1C); referred to as p57 in the rest of manuscript) is a newly identified interacting partner of CSN6 [116]. As p27, p57 is a potent natural inhibitor of several G1 CDKs. Overexpression of p57 was shown to completely arrest the cell cycle at G1 phase, whereas decreased levels of p57 were identified in many types of cancers [117]. The degradation of p57 is ubiquitination-dependent and is mediated by the SCF<sup>Skp2</sup> complex [118]. It was found that CSN6 is involved in p57 downregulation and results in an increase in p57 ubiquitination levels [116]. CSN6 interacts with both p57 and Skp2, which, in return, enhances Skp2-mediated p57 ubiquitination and as such results in a decrease in p57 expression levels. Importantly, increased expression levels of CSN6 in addition to decreased p57 expression levels correlate with poor human tumour survival [116], which is reminiscent to the observed negative correlation between CSN5 and p27 in various cancers, as described above.

### **1.3 CSN in different species:**

#### **1.3.1 The architecture of the CSN in different species**

The CSN complex was initially identified in *At*, as a negative regulator of the photomorphogenesis developmental, which occurs in response to light sensitivity in the team of Xing Wang Deng at the University of Yale (USA) by Ning Wei. Later on, during the course of the purification of proteasomes from mammalian red blood cell lysate, the CSN was further isolated [119]. Moreover, the complex was also purified from pig spleen [120]. The

presence of the CSN in both plants and mammals indicated a function that lies beyond the control of light-mediated plant signal transduction. Further discoveries of the CSN complex were later reported in a variety of eukaryotic organisms, from unicellular to chordates, including *Saccharomyces cerevisiae* (*Sc*), *Schizosaccharomyces pombe* (*Sp*), *Aspergillus nidulans* (*An*), *Neurospora crassa* (*Nc*), *Caenorhabditis elegans* (*Ce*), *Drosophila melanogaster* (*Dm*) and *Hs*. To date, the CSN complex has been identified in mammals, plants, and yeast and functionally has been shown to physically associate with SCFs [47, 121].

The subunit composition of the CSN complex, responsible for the structure and function of the complex displays a considerable variation in different species (Table. 1). The CSN complex in mammals, plants and insects typically comprises eight subunits, six of which are PCI domain-containing and two are MPN domain-containing subunits [122]. In addition the archaetypal CSN shares structural features with the 19S proteasome lid or the eIF3 complex. Yeast CSN appears to be more divergent in its composition, potentially lacking two subunits.

**Table 1-1. CSN and 19S proteasome lid subunit composition in different organisms.**

\* Subunit dually present in both the CSN and the 19S proteasome lid

Subunit	<i>Sc</i>	<i>Sp</i>	<i>Ce</i>	<i>Dm</i>	<i>Hs</i>	<i>At</i>
PCI	Csn11 Rpn7	Csn1 Rpn7	CSN1 RPN-7	CSN1 Rpn7	CSN1 Rpn7	CSN1/FUS6 RPN7
PCI	Csn10 Rpn6	Csn2 Rpn6	CSN2 RPN-6	CSN2 Rpn6	CSN2 Rpn6	CSN2/FUS12 RPN6
PCI	- Rpn3	Csn3 Rpn3	CSN3 RPN-3	CSN3 Rpn3	CSN3 Rpn3	CSN3/FUS11 RPN3
PCI	Rpn5* Rpn5	Csn4 Rpn5	CSN4 RPN-5	CSN4 Rpn5	CSN4 Rpn5	CSN4/COP8 RPN5
MPN+	Csn5 Rpn11	Csn5 Rpn11	CSN5 RPN-11	CSN5 Rpn11	CSN5 Poh1	CSN5 RPN11
MPN-	Csi1 Rpn8	- Rpn8	CSN6 RPN-8	CSN6 Rpn8	CSN6 Mov34	CSN6 RPN8
PCI	Csn9 Rpn9	Csn7 Rpn9	CIF-1 RPN-9	CSN7 Rpn9	CSN7 Rpn9	CSN7/FUS5 RPN9
PCI	- Rpn12	- Rpn12	- RPN-12	CSN8 Rpn12	CSN8 Rpn12	CSN8/COP9 RPN12

### 1.3.2 The mammalian CSN complex

Several mammalian CSN subunits were originally identified as independent interactors of various signalling proteins. Particularly, the CSN5 subunit was initially reported as Jab1 in a potentially CSN-independent form that reflects its ability to directly interact with and stimulates c-JUN transcriptional activity [123]. Further studies found CSN5 interacting with various intracellular regulators where examples include, p27 [107], as described above, the leucocyte antigen receptor [124], various transcription factors and tumour suppressors like

p53 [88]. In mammalian cells, besides the holo-CSN complex containing eight subunits, the presence of smaller CSN subcomplexes was identified by analytical gel filtration analysis of cell lysate [125-128]. Although these sub-complexes could be intermediates along the CSN assembly/disassembly pathways, some have been shown to hold specific biological roles [127]. However, the molecular identity and functions of CSN subcomplexes remain to be clarified, although they may contribute to the multifunctionality of the mammalian CSN subunits. Interestingly, in mammalian cells, the CSN7 subunit exists as an isoform, CSN7a and CSN7b, which could define different types of CSN complexes with different functions, although this has not been investigated to date [129]. Proteomic analysis performed on the mammalian CSN complex demonstrates that, as expected by the typical composition of the CSN (one copy of each of the eight CSN subunits), CSN7b is not part of complexes containing CSN7a, indicating that CSN7a and CSN7b are indeed components of different CSN complexes [130].

To date, probably through its regulatory role of CRLs, the mammalian CSN was shown to be involved in a range of biological responses. However, in contrast to the CSN identified in the plant kingdom, the mammalian CSN is closely involved in cell proliferation and maintenance. Studies reported that knockout (KO) of various *CSN* genes, including those of CSN2, CSN3, CSN5 and CSN8, lead to lethality at an early embryonic development stage in mice, mainly due to defects in cell proliferation and survival [110, 131-133]. In addition, for the *CSN2* gene KO experiments, the poor levels of substrate proteolysis in these KO mice, lead to elevated levels of p53 and cyclin E cell cycle regulators [131]. Moreover, *CSN* knockouts in mouse T cell lineages show reduced T cell receptor activation and proliferation, whilst *CSN8* KO cells demonstrate higher levels of neddylated Cul2 and of cyclin E [134].

### 1.3.3 The CSN in plants

In *At*, and as in all identified higher organisms, the CSN complex comprises eight subunits, six of which were genetically identified as pleiotropic, constitutively photomorphogenic/de-etiolated/fusca (*COP/DET/FUS*) mutants [135]. Although this class of mutants showed survival through embryogenesis, their growth is severely limited, defects that arrest growth at the seedling stage was identified and simultaneously irregular gene expression patterns was detected. Interestingly, in *At*, the CSN5 and CSN6 subunits are both encoded by two homologous genes, *CSN5a* and *CSN5b* or *CSN6a* and *CSN6b*, respectively [125]. For both *CSN5* and *CSN6*, mutants generated that lack one of the homologous genes are viable, highlighting their functional redundancy. However, KO of both homologous genes results in lethality, showing the common phenotype obtained in multicellular organisms [136-138]. Besides *At*, several other plant species also possess two *CSN5* and *CSN6* genes, however duplication of these genes has only been observed in the plant kingdom [139]. The functional reason for these gene duplications in the *At*CSN system remains undetermined, although in general, gene duplications are commonly present to produce functional diversification or specialisation.



#### 1.3.4 CSN in yeast and fungi

Functional studies in higher eukaryote organisms, summarised in the section above, have shown that the complex is essential for development and that defects in the CSN are responsible for embryonic death. The CSN complex has been identified in various fungal species including *Nc*, *Sp*, and *Sc*, showing significant differences in the composition of the complex with fewer than eight subunits in comparison to the canonical eight-subunit complex observed in higher eukaryote organisms.

Purification of the CSN complex from *Nc* identified a complex composed of subunits 1-7, whilst the CSN8 subunit was found to be lacking not only in the purification experiment but also in the genome sequence [140, 141]. This study further suggested the implication of the CSN in *Nc* circadian clock and showed that the CSN is required for the stability of SCF complexes in *Nc* [140, 141].

The CSN in *Sc* is largely divergent containing one MPN domain-containing protein which is the ortholog of the CSN5 subunit, Rri1, and four PCI subunits, CSN9, CSN10, CSN11 and CSN12, which show only few similarities to subunits present in higher eukaryotes [142]. An additional interactor, Csi1, has been found to associate with the 5-subunit CSN-like complex found in *Sc* [142]. Interestingly, the Csi1 displays significant homology with the C-terminal domain of the canonical CSN6 subunit found in mammalian CSNs, but lack the N-terminal MPN domain [143]. It was further proposed that the Csi1 comprises the functional equivalent of the canonical CSN6 [143].

Further disparities have been reported in the *Sp* CSN complex, which is composed of subunits CSN1-5 and CSN7a, whilst the subunits CSN6 and CSN8 as missing and neither subunit appear to contain any recognizable homolog [144]. In these organisms, the CSN complex is involved in regulation of diverse pathways [140, 145, 146]. Interestingly, *csn1* and *csn2* mutants in *Sp* identified a delay in S phase progression and showed hypersensitivity to UV - and  $\gamma$ -radiation [146], whilst deletions of *csn4* and *csn5* in this organism revealed that both subunits are required for deneddylation of the *Sp* Cullin Pcu1. However, neither *csn4* nor *csn5* showed to share common phenotypes with *csn1* and *csn2* mutants [144]. These findings indicate a large variation in the CSN complex composition and function between fungi and higher eukaryotes. However, more recently, it was shown that the mold *An* also contained the eight-subunit CSN complex and most interestingly defects showed to result in the abolishment of the potential to form fruit bodies, producing a phenotype reminiscent to that of higher organisms [147]. The fungal kingdom is thus diverse in itself, as well as showing the largest variation.

## 1.4 CSN isopeptidase and other isopeptidases:

### 1.4.1 Isopeptidases

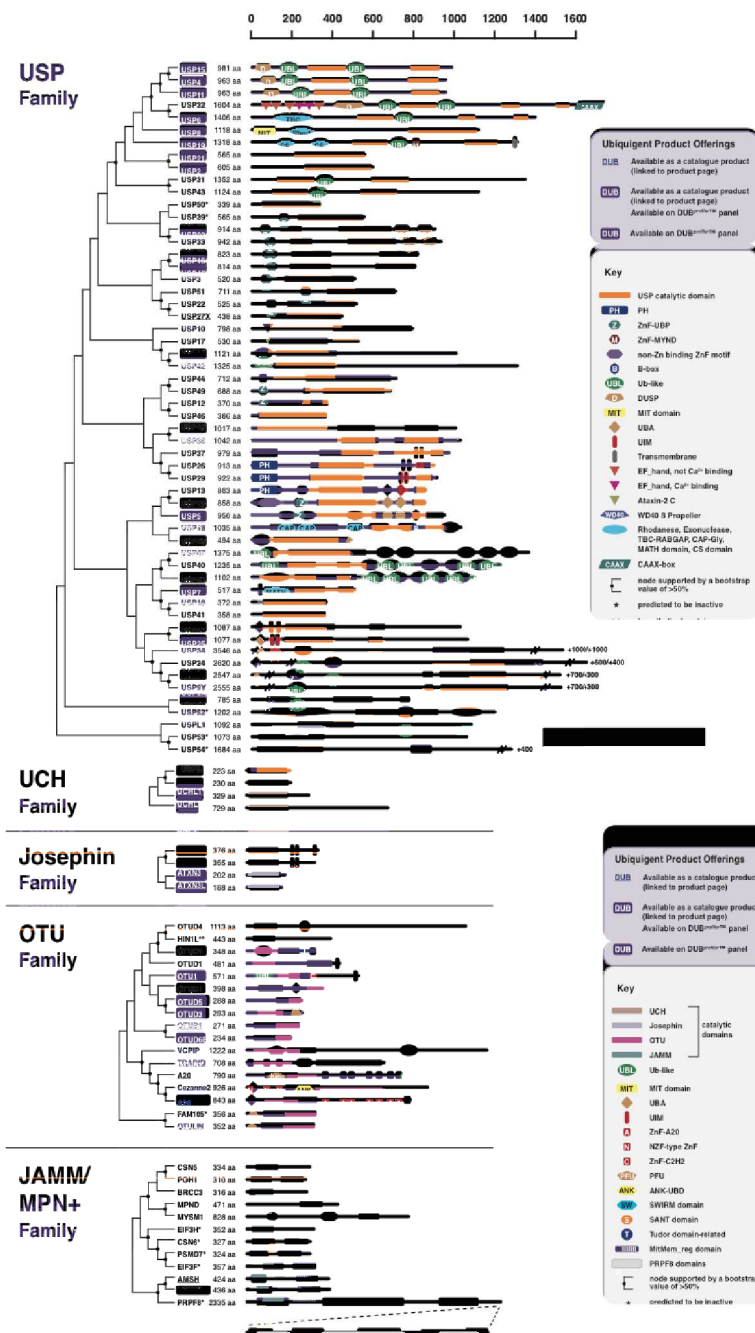
As for other post-translational modifications, the conjugation of Ub and Ubl molecules to their substrate is a reversible process [148]. The isopeptidase family classifies a diverse population of proteolytic enzymes involved in counteracting these modifications. The generic name of these enzymes is generally defined by their substrate. This family of proteins counts DUBs, which in principle are specifically implicated in the cleavage of Ub linkages. Likewise, enzymes that hydrolyse isopeptide bonds implicating Nedd8, ISG15 or SUMO are referred to as deNEDDylating, deISGylating, or deSUMOylating enzymes, respectively. In general terms, they can also be viewed as agonists of E3 ligases, but they can also remodel Ub or Ubl chains.

Based on their catalytic mechanism, proteases have been divided into five classes, comprising of aspartic, cysteine, metallo, serine and threonine proteases [149]. The cysteine and metallo protease classes contain DUBs [149]. Genomic studies have identified 95 human genes encoding for putative DUBs, among which the majority are cysteine proteases [150]. Whilst all metalloprotease domain DUBs hold a JAMM/MPN<sup>+</sup> domain, cysteine proteases can be further subdivided according to their Ub protease domains into the UCHs, Ub-specific proteases (USPs), ovarian-tumor (OTU) domain proteases, and Machado-Joseph domain proteases (Fig. 1-5) [149]. In this section, we mainly focus on the metalloprotease DUBs and specifically on members of this subfamily, which within our work we highly referenced, in order to provide functional understanding of our system. However, further functional and structural properties of DUBs are addressed in Chapter 3.

The large diversity observed in DUBs is suggestive of considerable substrate specificity [151]. Linkage-specific cleavage, sensitivity for poly-Ub chains and variation in efficiency ensure another level of specificity of isopeptidases [152, 153]. Examples include the isopeptidases, USP8 and USP14 which cleave K48 but not K63 polymers of Ub [149], whilst K63-linked Ub is hydrolysed by the cylindromatosis gene product (CYLD) [154]. Moreover, some DUBs are able to deconjugate both Ub and Ubls from their substrate [155]. This is the case for USP21 and UCH-L3 that can cleave Ub, as well as Nedd8 [156]. Although these enzymes show to contain low specificity *in vitro*, their ability to cleave both modifiers suggests a crosstalk between levels of DUBs. It has further been shown that the subcellular localisation and other scaffolding partners of DUBs influence their specificity and potentially regulate their activation [151, 157]. A study investigating the DUB interaction landscape using proteomic analysis investigated 75 of the 95 DUBs encoded in the human genome and identified 774 unique high-confidence candidate interacting proteins. This work led to the interesting observation indicating that a good number of DUBs are associated with multi-subunit complexes and further supports the hypothesis that the incorporation of DUBs with other interacting partners is required for their activation and specificity [150]. Indeed it has been shown to be the case for the catalytic subunits of the 19S proteasome lid, the CSN complex,



and BRAC1-RAP80 and BRISC complexes, Rpn11, CSN5, and BRCC36, respectively (Fig. 1-6) for JAMM DUBs [47, 75, 150, 158-160].



**Figure 1-5. The human DUBomes.**  
Adapted from Ubiquigent ([www.http://ubiquigent.com](http://ubiquigent.com)).

From a functional point of view, the isopeptidase activity of DUBs serves for various purposes [148]. Since most Ub and Ubls are encoded as a precursor, there are DUBs implicated in their processing that is required for their function as post-translational modifiers. Second, DUBs are involved in the deconjugation of Ub or UBLs from their substrate influencing protein stability or their involvement in signalling events. These can be described as DUBs that operate as negative regulators of the post-translational modifications as they antagonise the E3 ligase activities. Additionally, some DUBs function to cleave Ub from proteins that are tagged for degradation, required for efficient substrate degradation within the UPS and also for recycling the pool of unconjugated Ub [148]. Lastly, some DUBs are also involved in the remodelling of Ub and Ubl chains, prior for example to substrate degradation [159, 161, 162]. In the following sections, a detailed overview of examples of DUBs is given, focusing on the function of their isopeptidase activity. Whilst the activity of some DUBs requires their incorporation within larger complexes, this is not the case for all (Fig. 1-5).

#### 1.4.2 The CSN Isopeptidase

The major function of the CSN encompasses the proteolytic cleavage of the isopeptide bond between Nedd8 and the C-terminal conserved lysine residue of the Cullin subunit [47, 70, 163]. Neddylation of Cullins is a reversible reaction counteracted by the reaction of deneddylation that removes the Nedd8 moiety by the isopeptidase activity of the CSN [47]. The isopeptidase activity of the CSN complex is centred on its CSN5 subunit containing a JAMM/MPN<sup>+</sup> domain. Specific mutation performed on CSN5 that abolish its enzymatic activity results in accumulation of neddylated Cul1 in *Sp* [75], *At* [137, 138] and *Dm* [164] and neddylated Cul3 in *Ce* [165]. The effect on other Cullins has not been addressed. Moreover, in plants, loss of CSN activity as a result of deletion of various subunits was reported to lead to constitutive photomorphogenic phenotype. Such experiments in *Ce* resulted in sterility, whilst were found lethal in *Dm* and *Mm* [132, 166, 167].

As reported for other JAMM/MPN<sup>+</sup> proteins, the CSN5 subunit alone does not possess *in vitro* any isopeptidase activity and requires to be recruited within the CSN complex to display its deneddylase activity [70, 159, 168]. Deneddylation performed by the CSN is strictly conserved among all eukaryotes, although, as previously discussed, the subunit composition of CSN complexes differ between lower and higher eukaryotic species. Given the crucial role of the CSN5 subunit in the deneddylase activity of the CSN, its ability to associate with a range of cell regulators, the correlation of its expression level with cell proliferation, all suggest that the dynamics of CSN5 association is likely to be highly regulated. The lack of structural information available regarding the auto-inhibited nature of CSN5 in the stand-alone state challenged, until recently, the understanding of the CSN activity regulation.

### 1.4.3 The 19S proteasome lid isopeptidase

In cells, Ub signalling is essential and is implicated in diverse functions. It is thus inevitable that a highly controlled mechanism is required to prevent undesired deubiquitination of proteins by the proteasome. Within the proteasome lid, the enzyme responsible for removing the poly-Ub chain is Rpn11 (it can also be known as POH1 in *Hs*) [159, 161]. It is therefore entirely coherent that the DUB activity of Rpn11 is inhibited, until the lid is incorporated into the regulatory particle. Once bound to the 26S proteasome, via an Ub receptor, cleavage of the isopeptide bond between the protein substrate and the poly-Ub chain is required prior to subsequent substrate translocation through the axial pore of the proteolytic 20S chamber. Rpn11 cleaves the poly-Ub chain at the proximal Ub to permit subsequent fast degradation and recycles Ub for further use [161]. Failure to remove the poly-Ub chain from the substrate protein is thought to obstruct the entry of the substrate through the pore of the proteasome. This would result in a delay in the translocation of the substrate due to steric hindrance by the poly-Ub chain and additionally accumulate poly-Ub [169]. It has thus been shown that deubiquitination of the substrate by Rpn11 is required for efficient substrate proteolysis. Intriguingly, the DUB activity of Rpn11 is found to be ATP-dependent although it lacks an ATPase domain signature. As such it is suggested that the DUB activity is coupled to protein unfolding performed by the base ATPases neighbouring the lid [159, 170-172].

EM reconstruction studies revealed the six PCI domain-containing proteins of the lid assemble in a horseshoe-shaped arrangement that is suggested to act as a scaffold to stabilize the lid via substantial lateral interactions between the C-termini of PCI domain subunits [173-175]. The MPN domain-containing proteins, Rpn11 and Rpn8 (it can also be known as MOV34 in *Hs*), form a heterodimer and accommodate above the opening of the AAA-ATPase ring and incorporate into the scaffold formed by the PCI domain-containing subunits [175]. As observed for various macromolecular assemblies, the catalytic activity of the 19S proteasome lid is only imparted upon complete maturation of the complex [159].

The 19S proteasome lid and more specifically the DUB activity of the Rpn11 subunit is regularly referred to within our work. The common structural and architectural relevance with the CSN and CSN5 makes it suitable for comparison and for complementing our work. Initially, structural information on the MPN domain-containing proteins of the proteasome lid was restricted to the crystal structure of *Hs*Rpn8 MPN domain in the stand-alone state [176]. However the recently report structure and biochemical study of the MPN heterodimer, Rpn11-Rpn8, expanded the structural characterisation of this family and allowed validation of our work and a fine comparison [158, 160]. This aspect of our work is developed in Chapter 3.

#### 1.4.4 The eIF3 isopeptidase

The eIF3 complex is functionally implicated in the assembly of translation initiation complexes. It commonly acts as a docking site for various eIFs that assemble on the 40S ribosomal subunit [177]. This multi-protein complex contains a molecular mass of 800 kDa and comprises of 13 different subunits, namely eIF3a to eIF3m, within which six are PCI domain-containing proteins (eIF3a, c, e, k, l, and m), two are MPN domain-containing proteins (eIF3f and h), whilst the remaining are eIF3-specific subunits [57]. To date, structural insight into the eIF3 complex organization was obtained by cryo-EM studies, which highlighted the direct interaction between the MPN domain-containing subunits [178]. The enzymatic activity of the eIF3 complex has not yet been clearly characterised, however its intriguing resemblance in terms of subunit composition and topology, with the 19S proteasome lid and the CSN, has suggested potential to harbour DUB activity. Conversely, the JAMM motif constituting the catalytic site of DUBs does not appear strictly conserved in the putative MPN+ subunit, eIF3f, of this complex. However, it has been demonstrated that the MPN domain of the eIF3f subunit exhibits DUB activity specific for the activation regulation of the transmembrane receptor Notch. This work was further validated by mutational work on the active site residues of this subunit [179].

#### 1.4.5 The BRCC36-containing isopeptidases

For JAMM/MPN+ domain-containing subunits that are part of a large multi-protein complex, a recurrent principle is that they are catalytically inactive when in their stand-alone state. Concerning CSN5 and Rpn11 that are part of the CSN complex and the 19S lid complex, respectively, their incorporation into their respective complexes is required for their isopeptidase activity. A third multi-subunit complex containing a JAMM/MPN+ subunit includes the eIF3 complex that has recently been reported to express isopeptidase activity [179]. Besides containing a catalytic MPN domain-containing subunit, these complexes also have a MPN- subunit and six PCI proteins. Pairs of MPN+/MPN- proteins appear to be often present in these assemblies, although this does not seem to be the case for BRCC36 (for BRCA1 (for Breast cancer type 1 susceptibility protein 1)/BRCA2 (for Breast cancer type 1 susceptibility protein 2)-containing complex subunit 36)-containing complexes described to date. BRCC36 is an MPN+ protein functioning as a DUB in two large, possibly related, DNA repair complexes, the BRCA1-RAP80 (for Receptor-associated protein 80) complex and the BRCC36-containing isopeptidase complex (BRISC) [65, 159, 161]. In both complexes, BRCC36 possesses DUB activity, specifically for Lys63-linked poly-Ub chains [180-182].

The BRCA1-RAP80 is a five-member complex comprising of RAP80, BRCC36, BRCC45 (for BRCA1/BCRA2-containing complex subunit 45), Abraxas and MERIT40 (for Mediator of RAP80 interactions and targeting subunit of 40 kDa; also known as BRISC and BRCA1-A complex member 1). RAP80 contains an Ub-interacting motif that interacts with DNA double strand breaks and specifically recognize Lys63-linked chains [183, 184]. Together with the

isopeptidase activity of BRCC36, this complex is implicated in Lys63-linked Ub synthesis and hydrolysis at DNA double strand breaks thus regulating the recruitment and maintenance of DNA repair factors [181, 185].

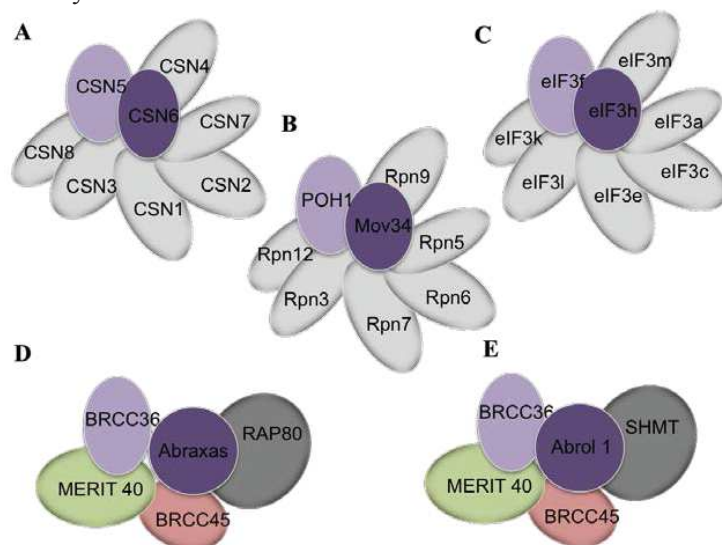
Recently, the BRISC complex was identified and share three subunits in common with BRCA1-RAP80 including, BRCC36, BRCC45, and MERIT40, whereas the fourth subunit namely Abrol (also known as KIAA0157) is a paralog of the Abraxas subunit from the BRCA1-RAP80 complex [182]. The DUB activity of BRCC36 in these complexes was individually characterised revealing that BRCC36 activity requires different interactions within the context of each protein complex. Within the BRCA1-RAP80 complex, the interactions of Abraxas and BRCC45 are essential for the DUB activity of BRCC36, whilst KIAA0157 is the only interaction necessary for BRCC36 DUB activity in the BRISC complex. Interestingly, both complexes require a protein-protein interaction with a MPN-domain-containing protein, that is the subunits Abraxas and KIAA0157, in BRCA1-RAP80 and in BRISC, respectively [168]. Regulation of DUB activity via protein-protein interactions was further validated by activity measurements of *HsRpn11* in the presence of the MPN-subunit of the proteasome lid, *HsRpn8*, revealing a catalytically active Rpn11 [160, 168]. These findings are suggestive of a common regulatory mechanism in this class of DUBs, where interactions between MPN+/MPN- proteins are essential for their isopeptidase activity.

#### 1.4.6 The AMSH-LP isopeptidase

The structure of AMSH-LP (associated molecule with the SH3 domain of STAM (AMSH)-like protease; also known as STAMBPL1) bound to a K63-linked di-Ub substrate and its biochemical characterisation led to the hypothesis that the mechanism of JAMM/MPN+ isopeptidases is presumed to be comparable to that of typical metalloproteases, in particular thermolysin [186]. However, the main difference is that proteases cleave peptide bonds, whilst JAMM/MPN+ enzymes are suited for the hydrolysis of the amide bond conjugating the C-terminus of distal Ub with an amino group on the proximal Ub moiety. In this study, two histidines, an aspartate and a water molecule are shown to coordinate the zinc ion. Mechanistically, it was deduced that an active site located glutamate residue activates the water molecule by accepting a proton from the catalytic water and thus transforms it to a more potent nucleophile. The resulting hydroxyl ion in return is able to attack the isopeptide bond of the substrate at the carbonyl carbon. A conserved serine present in the MPN+ motif stabilises the oxyanion formed. This mechanism results in the collapse of the transient tetrahedral bond of the substrate [186]. Sequence conservation of JAMM/MPN+ domain-containing proteins supports this mechanism for the other JAMM/MPN+ proteins discussed in this Chapter and following ones, including Rpn11, CSN5 and BRCC36. Moreover, the requirement of zinc for catalysis by these proteins has been demonstrated by inhibition of their activity in the presence of zinc chelators [70, 159].

The function of AMSH-LP DUBs has not yet been clearly understood. Their common structural characteristic with AMSH and high sequence identity suggests that it may carry out

a related function. These two proteins share 53% sequence identity and both contain a nuclear localisation signal and a JAMM/MPN+ domain [187]. Additionally both enzymes hydrolyse both Lys63- and Lys48-linked di-Ub substrate and are active in their stand-alone form, although domains outside of the DUB domain could affect the isopeptidase activity. The main difference between the two lies on the AMSH ability to associate with the SH3 domains of STAM1 (signal transduction adaptor molecule 1) and to clathrin. Although SH3-like domains are also present in AMSH-LP, it does not interact with STAM, suggesting that AMSH and AMSH-LP may have differing cellular functions. The function of AMSH encompasses the regulation of receptor endocytosis and contributes to the dynamics of receptor trafficking by deubiquitinating receptors tagged with a poly-Ub chain, thereby antagonising E3 ligase activity.



**Figure 1-6. Schematic representation of MPN+/MPN- domain-containing multi-protein complexes.**

Although MPN+ proteins are not all found in large protein complexes (including AMSH, AMSH-LP, MPND, MYSM1) here we focus on those that do. Additionally the MPN- protein, Prp8 also forms a complex with other protein, but lacks the MPN+ protein. In this schematic, color codes are used to characterise the domains present. JAMM/MPN+ subunits are in light purple, MPN- subunits in dark purple and PCI domain subunits in grey. The additional components present in the BRCC36-containing complexes are colored as indicated in the image. **A.** The human CSN complex. The complex contains the two MPN domain-containing proteins (CSN5 and CSN6) and six PCI subunits. **B.** The human 19S proteasome lid complex containing the two MPN subunits (*HsRpn11* (POH1) and *HsRpn8* (MOV34) that are the human orthologs of *ScRpn11* and *ScRpn8*, respectively) and six additional PCI subunits. **C.** The human eIF3 core complex. The schematic represents the two MPN subunits (eIF3f and eIF3h) and six PCI subunits. Another five non-MPN or PCI subunits are also present in the holo-eIF3 complex, including eIF3b, -d, -g, -i, -j (non represented here for clarity). **D.** The core of the BRCA1-RAP80 complex, the two MPN subunits (BRCC36 and Abraxas) are indicated and decorated with three other proteins that together form the core structure. **E.** The BRISC composed of the two MPN subunits (BRCC36 and KIAA0157/Abrol) and further three other subunits.

In our work, we use both AMSH and AMSH-LP as model systems to gain further insight on how MPN+ domain-containing proteins reach a catalytically competent form and more specifically addressing how the inactive stand alone CSN5 subunit is in an active conformation when incorporated into the CSN complex. Given that AMSH-LP bound di-

Ub and AMSH bound di-Ub and Ub [188] are currently the sole MPN+/substrate complex structure available, we employ this system to understand substrate recruitment in the case of CSN5. This work is further described in Chapter 3.



## Chapter 2: Materials and Protocols

### 2.1 Protein production:

#### 2.1.1 CSN5 Construct design, cloning, expression and purification

cDNA coding sequence for full-length (FL; 1-334) and a truncated version of CSN5 cloned into the pGEX-6P1 vector (GE Healthcare) were used to recombinantly express CSN5. Site directed mutagenesis was performed using the QuikChange Lightning Site-Directed mutagenesis kit (Stratagene) and point mutation oligonucleotides (Eurogentec). All constructs and mutations were checked by DNA sequencing (Beckman Coulter Genomics).

Expression and purification of CSN5 (fragment 1-257) and mutant forms in BL21pLysS *E. coli* cells (Novagen®) were performed using standard conditions. Induced bacterial cells were resuspended in the purification buffer (20 mM Na-2-(N-morpholino)ethanesulfonic acid (Na-MES), pH 6.5, 100 mM NaCl, 0.002% monothioglycerol (MTG)) supplemented with EDTA-free protease inhibitor cocktail (Roche) and were lysed by sonication. Cell lysate was clarified by centrifugation and applied onto a gravity-flow Glutathione Sepharose 4B column (GE Healthcare). Glutathione *S*-transferase (GST)-tagged CSN5 was eluted by 20 mM reduced glutathione in the purification buffer and was cleaved overnight at 4 °C by GST-3C protease. The sample was concentrated and loaded onto a Superdex 75 gel filtration column (GE Healthcare). A final polishing step was used to separate CSN5 from contaminating GST. The resulting pure CSN5 was concentrated to 10 mg mL<sup>-1</sup> and stored at -80°C until further use. Protein concentrations were measured with a Nanodrop (ThermoScientific) at 280 nm using their theoretical extinction coefficient.

#### 2.1.2 CSN6 Construct design, cloning, expression and purification

The human CSN6 cDNA was subcloned in pGEX-6P1 (Novagen®). Soluble heterologous expression of glutathione *S*-transferase (GST)-tagged full-length (FL) CSN6 in bacteria proved challenging. As a result, three C-terminal truncated forms of CSN6 were designed and the fragment comprising the residues 31 to 211 was selected for further work. Expression of CSN6 and its variants was performed in *E. coli* BL21pLysS cells (Novagen®) in LB or <sup>15</sup>NH<sub>4</sub>Cl-supplemented M9 minimal expression medium, as required for NMR experiments. Cells were induced with 0.3 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) overnight at 18°C. Harvested cells were resuspended and lysed in purification buffer (20 mM Tris pH7.5, 150 mM NaCl, 0.01% monothioglycerol (MTG)), supplemented with EDTA-free protease inhibitor cocktail (Roche) and later sonicated. Following centrifugation, clarified cell lysate was applied on a gravity-flow glutathione Sepharose 4B column (GE Healthcare). GST-tagged CSN6 was eluted with 20 mM reduced glutathione in purification buffer. Cleavage of the fusion tag was obtained by incubation of the eluant, overnight, at 4°C, with GST-3C



protease (1/50 w/w). The sample was concentrated and applied onto a Superdex 75 gel filtration column (GE Healthcare), pre-equilibrated in the purification buffer supplemented with 5 mM benzamidine to avoid C-terminal proteolysis. The eluted fractions corresponding to CSN6 were further purified on a glutathione Sepharose 4B column to remove contaminating GST and GST-tagged proteins. Purified CSN6 was concentrated to 8 mg mL<sup>-1</sup> and stored at -80°C until further use.

### **2.1.3 Construct design, cloning, expression and purification of other proteins used in this project**

Production of NEDP1 (Nedd8-specific protease 1; also known as DEN1 for Deneddylase-1 or SENP8 for SUMO/Sentrin Specific Peptidase Family Member 8) from the pGEX-2TK-nedp1 vector (kind gift from Dr. D. Xiroadimas) was carried out following the protocol described in [189]. The pOPIN-E-pro-nedd8 vector (kind gift from Dr. M. Banfield) was used to express pro-Nedd8 following the procedure described in [190].

The Cullin1/Rbx1 complex was obtained by co-expressing the set of the two ‘split-an-express’ plasmids (pAL-hCul1-NTD, pCool-mRbx1-His-hCul1-CTP) kindly provided by N. Zheng (University of Washington) and distributed by Addgene, following the procedure described in [191]. GST-Rbx1/Cul1-CTP/Cul1-NTD was purified from a 0.9 L bacterial culture using a glutathione sepharose column. After thrombin cleavage, the protein solution was subjected to a separation on a Mono S 5/50 column. Fractions corresponding to the Rbx1/Cul1-CTP/Cul1-NTD were pooled, concentrated and stored at -80°C until further use.

In Chapter 4, the Cullin1-Nedd8/Rbx1 and CSN were kindly provided by Prof. M. Peter and Dr. R. I. Enchev (ETH Zürich), following procedures described in [46].

## **2.2 Enzymatic assays:**

### **2.2.1 Isopeptidase assays on synthetic substrates**

For the isopeptidase assay using LRGG-AMC substrate performed in Chapter 3, GST-tagged CSN5 1-257 protein and different mutants were diluted to 0.2 µg µL<sup>-1</sup> in reaction buffer (40 mM Tris-HCl pH8.5, 5% glycerol, 1 mM DTT), in the presence of 250 µM LRGG-AMC. For the isopeptidase assay using Nedd8-AMC substrate, CSN5 1-257 WT protein and different variants (R106T, R106A, R106G, R106P, E76A, E76A/R106T) prepared in the same conditions were diluted to 0.2 µg µL<sup>-1</sup> immediately before in the reaction buffer (40 mM Tris-HCl pH8.5, 1 mM DTT). The CSN complex purified from erythrocytes and purchased from Enzo Life Sciences was used at 0.01 µg µL<sup>-1</sup>. The reactions setup on ice were started by the addition of the substrate (2 µM Nedd8-AMC) to the reaction mixture and followed at 28°C. Isopeptidase assays were monitored in duplicate in a 96-well fluorescence plate on a Tecan

Sapphire, by following the increase of fluorescence intensity ( $\lambda_{\text{excitation}}=380$  nm;  $\lambda_{\text{emission}}=460$  nm), i.e. the hydrolysis of the isopeptide bond between LRGG/Nedd8 and AMC.

For experiments performed in Chapter 4, initial rate measurements were carried out at 37°C by following the increase of fluorescence intensity ( $\lambda_{\text{excitation}} = 380$  nm;  $\lambda_{\text{emission}} = 460$  nm), in non-binding surface, flat-bottom, black 384-well plates (Corning), in a reaction volume of 50  $\mu\text{L}$  on a Tecan safire<sup>2</sup> plate reader. For the LRGG-AMC substrate, the following experimental conditions were used: reaction buffer (50 mM Tris-HCl pH7.5, 50 mM NaCl), LRGG-AMC concentrations (25 to 400  $\mu\text{M}$  in 50 mM Tris-HCl pH7.5, 50 mM NaCl, 1.25% DMSO), 3.4  $\mu\text{M}$  CSN5/CSN6 binary complex (5.9  $\mu\text{M}$  CSN5 and 10.0  $\mu\text{M}$  CSN6 mixed prior to reaction), 3.4  $\mu\text{M}$  NEDP1. As saturating concentrations of LRGG-AMC were not reached, the value of the pseudo-second order rate constant ( $k_2$ ), corresponding to  $k_{\text{cat}}/K_M$  value was obtained by linear fitting of initial rate data. For the Nedd8-AMC substrate, the following experimental conditions were used: reaction buffer (50 mM Tris-HCl pH7.5, 50 mM NaCl), Nedd8-AMC concentrations (0.5 to 6.0  $\mu\text{M}$  in 50 mM Tris-HCl pH7.5, 50 mM NaCl), 4 nM CSN5/CSN6 binary complex (100 nM CSN5 and 169 nM CSN6 were mixed prior to reaction and the concentration of the binary complex was calculated using the value of the  $K_D$  determined by ITC), 200 nM CSN5 alone. When saturating concentrations of Nedd8-AMC were observed, the initial rate data were fitted to the Michaelis-Menten equation using least square analysis to determine  $k_{\text{cat}}$  and  $K_M$  in the Prism software (GraphPad). When saturating concentrations of Nedd8-AMC were not observed, the value of  $k_2$ , corresponding to  $k_{\text{cat}}/K_M$  value was obtained by linear fitting of initial rate data.

### 2.2.2 Pro-Nedd8 processing assays

For this assay, CSN5 (WT or variants), alone or the CSN5/CSN6 complex was diluted to 900 nM (for CSN5 alone) or to 4 nM (for the dimer: (100 nM CSN5 and 169 nM CSN6 were mixed prior to reaction and the concentration of the binary complex was calculated using the value of the  $K_D$  determined by ITC). Using the measured affinity between CSN5 and CSN6, each subunit was added in excess to reach a final binary complex concentration of 4 nM.) in reaction buffer (20 mM Tris-HCl pH7.5, 50 mM NaCl) in the presence of 1  $\mu\text{M}$  pro-Nedd8 WT or variants. A time course at 37°C was performed in independent duplicates for each reaction. The reactions were stopped in SDS sample buffer and analysed on Tris-tricine gel stained by Instant Blue solution (Expedeon, UK). Bands of processed pro-Nedd8 were detected and quantified using Carestream Molecular Imaging instrument connected to the Gel Logic 2200PRO software (Equilab, Whitestone, NY).

### 2.2.3 Deneddylation assays

The WT CSN5 protein and its R106T variant in the presence of CSN6, and recombinant CSN, were diluted to 4 nM in reaction buffer (20 mM Tris-HCl pH7.5, 50 mM NaCl) and incubated at 37°C in the presence of 1  $\mu\text{M}$  Cullin1-Nedd8/Rbx1. A time course (0 to 120 minutes)

was recorded for each reaction in independent duplicates. One time point was performed for CSN5R106T (900 nM) for 2 hours incubation time. Samples were analysed as described for pro-Nedd8 processing.

## 2.3 Biophysical characterisation:

### 2.2.1 Fluorescence anisotropy

To obtain quantitative information about the affinity of CSN5 to its substrate and to understand the function of CSN6 in activating CSN5, fluorescence anisotropy measurements were performed at 28°C, using a TECAN safire<sup>2</sup> plate reader (Tecan).

For these experiments, Nedd8 (commercially available from BostonBiochem®) was tagged with Alexa488. Nedd8 at 160 µM was incubated with 50 µg of Alexa-488 dissolved in DMSO at room temperature on a rotator for 3 hours. The reaction was quenched with the addition of 2 mM Tris-HCl, pH7.5. Free dye from the sample was removed by loading the sample onto a 2 ml Zeba Spin Desalting column (Thermo Scientific). The labelling was checked by running a 12% Tris-glycine SDS-page gel and visualising by fluorescence and by measuring absorption to obtain the ratio of Alexa-488-Nedd8: free Alexa-488. A labelling efficiency of 34% was obtained.

The fluorescence anisotropy experiments were performed in 20 mM Hepes pH 7.5 and 75 mM NaCl. Alexa-488-Nedd8 was used as the labelled protein, and the fluorescence was measured using excitation and emission wavelengths of 470 nm and 530 nm, respectively. Varying concentrations (0-600 µM) of CSN5 and CSN5 R106T were incubated in the wells containing a constant concentration of 4nM Alexa-488-Nedd8. The same procedure was followed in the presence of CSN6 where a 1:1 ratio of CSN5/CSN6 was used to make the complex concentrations indicated. For each data point the measurement was repeated in triplicates and the recorded data corresponds to the average anisotropy. The data obtained were plotted and analysed using the IgorPro software. The Hill Equation (1):

$$f(X) = rB + (rAB - rB) * X^n / (K_D^n + X^n) \quad (1)$$

was used for fitting the data in order to be able to deduct an estimate of the apparent equilibrium dissociation constant ( $K_D$ ) of each complex and the apparent cooperativity ( $n$ ) between the protein and its ligand. The equation represented here uses the minimum ( $rB$ ) and the maximum ( $rAB$ ) signals as normalisation factors representing the lower and upper asymptote of the titration. As there exists no evidence of cooperativity, the binding was considered as not highly cooperative, and in all cases, the data were fit with values of cooperativity equal to 1.

### 2.2.2 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) experiments were carried out on a MicroCal ITC<sub>200</sub> microcalorimeter (GE Healthcare, Piscataway, NJ) at 20°C. The protein samples were all buffer exchanged using PD-10 desalting column (GE Healthcare) into the ITC buffer (20 mM Na MES pH6.5, 75 mM NaCl). Protein concentrations were measured using a NanoDrop 1000 spectrophotometer (Thermo Scientific, Wilmington, DE), with the following extinction coefficients: 18,450 M<sup>-1</sup>cm<sup>-1</sup> and 56,840 M<sup>-1</sup>cm<sup>-1</sup> for CSN6 and CSN5, respectively. Titration of CSN5 (WT and variants; 100 µM) in the cell (200 µL) was performed by sequential addition of CSN6 (WT and variants; 1 mM; 1.8 µL injection volume). Integrated raw ITC data were fitted to a one site nonlinear least squares fit model using the MicroCal Origin plugin as implemented in Origin 9.1 (Origin Lab) after the control experiments (titration of the ligand from the syringe into the buffer) were subtracted.

### 2.3.3 Crystallisation, data collection and structure determination of human CSN6.

Purified CSN6 at 8 mg mL<sup>-1</sup> was subjected to nanolitre crystallisation trials using commercial screening kits. Crystals grown at 18°C, in 0.2 M Tri-Sodium citrate, 0.1 M Bis-Tris propane pH6.5 and 20% PEG3350, were harvested, cryo-protected in the crystallisation solution supplemented with 10% glycerol and immediately flash-frozen in liquid nitrogen. A dataset was collected at a 1.76 Å resolution at the ESRF beamline ID29 and processed using MOSFLM and SCALA from CCP4 suite [192]. The structure was solved by molecular replacement using Phaser with Mov34 as the search model (PDB code 2O95). The initial model was built into the electron density map using ARP/wARP [193]. Further refinement procedure was carried out by alternate cycles of manual rebuilding and REFMAC program. Data collection and refinement statistics are compiled in Table 4-1.

### 2.3.4 Approaches used for CSN5-CSN6 crystallisation trials.

Microseeding. After obtaining micro crystals from the crystallisation screens performed for the CSN5-CSN6 complex, the seeding technique was used to grow crystals. The rationale behind this technique is that seed provides a template on which further molecules can assemble and given the proper environment the seed will enlarge into a crystal. The microcrystals observed in the hit in the Pro-complex screen (composition: 0.1 M MgCl<sub>2</sub>, 0.1 M HEPES pH7, 15% PEG 4000) were fished from the well and were added in a 10 µL of mother liquor. A bead (known as a Seed Bead) was added into the mixture which was further vortexed to ensure the crystal crushing. The resulting mixture containing nuclei was used as the protein sample to screen the Pro-complex kit, and further kits including JSCG+ and PEGs Suite. Limited amount of kits were screened due to the limited number of microcrystals initially obtained.

In situ limited proteolysis for crystallisation. Dynamic proteins and flexible regions within proteins make the crystallisation of these highly difficult. The presence of flexible regions in CSN5 encouraged us to employ *in situ* limited proteolysis to remove such regions. Prior to

performing the experiment, limited proteolysis was carried out. Using the limited proteolysis kit (implemented by Dr. François Hoh) we tested 12 different proteases. This procedure involved the use of each protease in three different dilutions, D1, 100, and D1000. Protein at 2  $\mu$ g was added to each dilution and then incubated at 37°C for 60 minutes. The samples were further run on a 15% Tris tricine SDS-PAGE gel, and the fragments obtained after proteolysis were analysed. This approach identified trypsin as an optimum protease to use for *in situ* limited proteolysis during crystallisation. The protein sample was incubated with the protease in all three dilutions and incubated at room temperature for prior to being used for screening.

### 2.3.5 Molecular dynamics simulations.

The A chain from the CSN5 (fragment 1-257) crystal structure was used as the initial structure for MD simulations on the WT protein and the R106 (T, G and P) variants. For the simulations run on CSN5 dimer A-B the oligomeric arrangement of the PDB file was used (PDB code: 4F7O), whilst the A-A' dimer was built in chimera [194] using crystallographic operators. The missing loop (residues 198–218) was built using the MODELLER program [195]. All the water molecules from the initial model were removed except the one bound to the catalytic zinc. Rather than using covalent bonds or harmonic restraints to maintain the zinc environment, we employed the cationic dummy atom approach [196], which imposes orientational constraints for the four zinc ligands (H138, H140, G151 and water) in the tetrahedral configuration. For the simulations run on AMSH-LP, the initial structure from the PDB code: 2ZNR was used and the dummy atom approach was also employed for the two Zn binding sites. The remaining protocol for running the MD simulations was maintained for all the simulations. The protonation state of the other ionisable side chains was set to their normal values at pH7. The resulting structure was surrounded by a periodic octahedral box of TIP3P water. For CSN5 monomer, this procedure resulted in a total of 4,029 protein atoms, including the zinc ion and the catalytic water molecule, solvated by 17,000–18,500 water molecules. All MD simulations were performed with the AMBER11 program [197] with the ff03 force field parameters [198] and the additional force field for the zinc atom environment [196]. Optimization and relaxation of solvent were initially performed by means of energy minimizations and MD simulations while keeping the solute atoms constrained to their initial positions with weak force constants.

After equilibration was established by gradually increasing the temperature from 100 to 300 K for 50 ps, the system was subjected to short (100-ps) MD simulations with decreasing constraints at a constant temperature of 300 K and a constant pressure of 1 bar. The 40-ns production run was conducted with constrained bond lengths involving hydrogen atoms and using the Shake algorithm [199], the Verlet integrator with a 2 fs time step for seven the calculation of forces and Langevin dynamics for temperature control. A cut-off radius of 9 Å was used to compute the non-bonded steric interactions. The electrostatic interactions were calculated with the particle-mesh Ewald method [200]. The missing counterions were substituted with a net-neutralizing plasma over the periodic box. The ptraj module in the AmberTools package [201] was used to extract data from trajectories and to analyze

structural and dynamic properties. All computations including the minimizations and the MD simulations were performed on a HP Z800 workstation equipped with two GPU Tesla C1060 and quad-core Xeon 2.4 GHz processors.

### 2.3.6 NMR.

Uniformly  $^{15}\text{N}$ -labeled CSN6 protein samples were dissolved in 50 mM Tris buffer, 150 mM NaCl (pH adjusted to 6.8) in 90%  $\text{H}_2\text{O}$ /10%  $\text{D}_2\text{O}$ . In all experiments, the  $^1\text{H}$  carrier was centred on the water resonance and a WATERGATE sequence [202, 203] was incorporated to suppress solvent resonances.

*Assignment of CSN6 (carried out by Dr. Y. Yang and Dr. A. Padilla).* Spectra were acquired with a 600  $\mu\text{M}$  CSN6 protein sample on a 700 MHz Avance Bruker spectrometer equipped with triple-resonance ( $^1\text{H}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}$ ) z-gradient cryo-probe at 298 K. NMR data were processed using GIFA (v 4.0) [204], and Topspin (v. 2.1). The 3D  $^1\text{H}$ - $^{15}\text{N}$ -NOESY-HSQC spectrum was analysed using strip-plots with manual reordering of the sequential stretches according to the main chain assignment strategy [205-207]. Side chain assignments were carried out using the  $^1\text{H}$ - $^{15}\text{N}$ -TOCSY-HSQC spectrum. The first two residues glycine and proline in the N-terminal tag were not assigned. For the remaining of the sequence, assignments were obtained for  $^{15}\text{N}$ , HN and Ha atoms (excluding proline residues, stretches G40-L42, L59-L60 and D71, H130 residues). The overall assignment completeness for  $^1\text{H}$  resonances was 79%.

*CSN6/CSN5 complex (carried out by Dr. Y. Yang and Dr. A. Padilla).* An HSQC spectrum at 308 K was acquired starting from a  $^{15}\text{N}$ -labeled CSN6 sample at 100  $\mu\text{M}$ . After addition of 200  $\mu\text{M}$  unlabeled CSN5 to the  $^{15}\text{N}$ -labeled CSN6 sample another HSQC was recorded.

### 2.3.7 Protein-protein docking and modelling.

CSN5 and CSN6 models were treated as rigid bodies and all six rotational and translational parameters were fully explored in the ZDOCK program using  $6^\circ$  sampling on Euler angles. A list of 30 ‘passive’ solvent-accessible residues of CSN6 excluded from the interface with CSN5 was derived from the analysis of non-shifted CSN6 peaks in HSQC experiments in the presence of CSN5 and incorporated in the filtering option of the ZDOCK procedure. Among the 54,000 models generated, the top 2,000 complexes, ranked according to ZDOCK were further analyzed. The ZDOCK scoring function combines pairwise shape complementarity, desolvation and electrostatics. The starting structures for human CSN5 and CSN6 proteins were the crystal structures solved at 2.5 Å and 1.8 Å resolution, respectively. Two docking pairs of protein models were used for docking simulations: the CSN5 (chain A, PDB code: 4F7O), with Ins-1 segment (residues 98-132) fitted to the corresponding ubiquitin-liganded AMSH-LP Ins-1 structure (PDB code: 2ZNV) using MODELLER9 program [208], combined with the 31-200 CSN6 fragment (deletion of the swapping C-terminal helix); and the same



CSN5 model with the N- and C-terminal appending segments removed (residues 2-51 and 232-257) combined with the fragment 31-200 of CSN6. In order to choose a near-native structure from these ZDOCK predicted sets, we filtered and re-ranked the 2000 hits using a dedicated procedure incorporating experimental constraints derived from biophysical and biochemical data. Additional information was also incorporated in this scoring procedure; primary sequence conservation data for both CSN5 and CSN6, total buried area in the dimer interface (calculated using the Naccess program (<http://www.bioinf.manchester.ac.uk/naccess/>), surface shape complementarity (obtained by SC program [192, 209]) and interface parameters derived from the PISA program [210]. CSN5, CSN6 and Nedd8 sequence conservation scores were evaluated using the PAT server [211]. Docking simulations were performed on two Intel Linux cluster platforms.

### **2.3.8 Analytical ultracentrifugation (Performed by Dr. C. Ebel, IBS, Grenoble)**

Sedimentation velocity experiments were performed using a Beckman XL-I analytical ultracentrifuge and an AN-50 TI rotor (Beckman Coulter). The experiments were carried out at 20 °C for CSN51-257 WT or L237QL240K variant at 333.0, 167.0, 67.0, 33.0, 17.0 and 3.3  $\mu$ M in 20 mM NaMES pH 6.5, 100 mM NaCl. A volume of 50  $\mu$ L (for the four most concentrated samples), 100  $\mu$ L or 400  $\mu$ L (for the most diluted one) was loaded into 1.5, 3.0 or 12 mm path cells and centrifuged at  $130,000 \times g$  (42,000 rpm). Scans were recorded every 22 min, overnight, by absorbance at 250 and/or 280 nm and by interference. The Sednterp software (freely available at <http://www.jphilo.mailway.com/>) was used to estimate the partial specific volume of the polypeptide chain,  $v = 0.734$  mL g<sup>-1</sup>, the solvent density,  $\rho = 1.009$  g mL<sup>-1</sup>, and the solvent viscosity,  $\eta = 1.02$  mPa.s, at 20 °C. Sedimentation profiles were analyzed by the size-distribution analysis of Sedfit. The isotherm of the mean sedimentation coefficients was fitted in term of a monomer-dimer equilibrium constant using the program Sedphat. Sedfit and Sedphat are freely available at <http://www.analyticalultracentrifugation.com>.

## **2.4 *In vivo* approaches:**

### **2.4.1 Co-immunoprecipitation experiments (Performed by Dr. F.-X. Claret, MD Anderson, Houston, USA)**

Cell culture medium were from Mediatech Inc (Mannassas, VA) and fetal bovine serum (FBS) were obtained from Gibco (Grand Island, NY). The antibodies used were His- tag (Cell Signaling Technology, Beverly, MA); Flag-tag and  $\beta$ -actin (Sigma-Aldrich, St. Louis, MO). The Lipofectamine Plus and Oligofectamine reagents were from Invitrogen (Carlsbad, CA). Western Lightning Chemiluminescence Plus reagent was from Thermo Scientific Pierce

(Rockford, IL). As described in (29), 293T cells were maintained in DMEM supplemented with 10% foetal bovine serum and 100 µg mL<sup>-1</sup> penicillin and streptomycin. Transfection was carried out using the Lipofectamine Plus reagent (Invitrogen). For the co-immunoprecipitation experiments, cells in the log-phase of growth were collected, washed twice in cold phosphate-buffered saline (PBS), and lysed as described previously (30). Cell lysates were incubated in RIPA buffer for 4 hours at 4 °C with either anti-Flag or anti-His antibodies. Proteins were separated by 12% SDS-PAGE, transferred to nitrocellulose membranes, and probed with anti-His, and anti-Flag antibodies. β-actin was used as the internal positive control for all immunoblots. Immunoreactive bands were detected using HRP-conjugated secondary antibodies with the Western Lightning Chemiluminescence Plus reagent. The protein levels were quantified using ImageJ software (National Institute of Health, Bethesda, MD, USA. <http://rsb.info.nih.gov/ij>).

#### **2.4.2 Growth of NEDP1 MEFs cells (Performed by Dr. D. Xirodimas, CRBM, Montpellier)**

U2OS cells and NEDP1 MEFs were grown in DMEM medium supplemented with 10% FCS (Foetal calf serum) and antibiotics. U2OS cells were plated in 6-well plates and transfected with either control or NEDP1 or CSN5 siRNAs (Pools of 4 individual siRNAs for each target from Dharmacon). 48 hours post-transfection, cells were harvested and lysed in NP-40 lysis buffer. NEDP1 MEFs were plated in 10cm dishes and were harvested when reached 80% confluency and lysed as above. Total protein was measured by Bradford assay and cell extracts were used either in Nedd8 processing experiments or for western blotting to monitor expression of NEDP1 and CSN5.

#### **2.4.3 Enzymatic assays performed on cell lysates**

To check the pro-Nedd8 processing activity in NEDP1 MEFs, in si-RNA cells and in control cells, initially all the cells were lysed. The thawed cells were resuspended in 200 µl of lysis buffer (50 mM Tris pH7.5, 150 mM NaCl, 1% NP-40 and 2 mM DTT) by pipetting. The cells were left to incubate on ice for 30 mins before centrifuging at 13,000 rpm for 15 minutes at 4 °C. The supernatant was then transferred and used for the enzymatic assays performed on pro-Nedd8 substrate. For the pro-Nedd8 assays using cell lysate, the amount of cell lysate added was adjusted according to the Bradford assay performed to use equal amounts of protein for each reaction. The cell lysates were spiked by 1 µM of pro-Nedd8 and the assay was run as described in section 2.2.2.



## Chapter 3: CSN5 isopeptidase activity

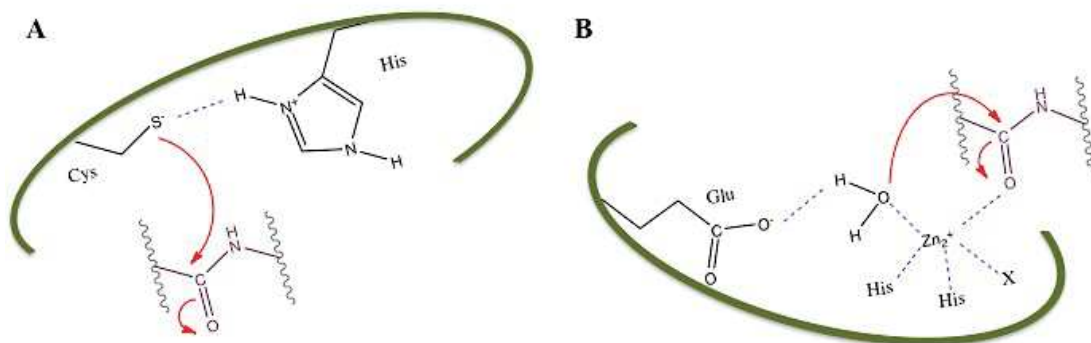
### 3.1 Introduction:

Gaining insights into the activity mechanism and the associated regulation by which the CSN promotes the cleavage of Nedd8 from Cul1 is being pursued and recent developments brought crucial elements. As mentioned in Chapter 1, CSN, through its activity in the UPS, is a critical player in a diverse range of cellular and developmental processes. The most studied function of the CSN complex is the regulation of CRLs through its deneddylating activity on Cullins, the scaffolding component of CRLs. CSN catalytic activity involves the hydrolysis of an isopeptide bond between a Nedd8 molecule and a specific C-terminal lysine residue of the Cullin scaffolding protein (K720 for human Cul1). The deneddylation of Cullins is thought to be a critical element for maintaining the stability and the continuous activity of CRLs *in vivo*, enabling spatial and time controlled of poly-ubiquitination of a large proportion of the ubiquitome [212]. The catalytic activity displayed by the CSN enables it to be classified as an isopeptidase, where it hydrolyses an isopeptide bond specifically between Nedd8 and an exposed lysine residue at the C-terminal of Cullins.

In cells, Ub and most UbIs (including Nedd8 and SUMO) are generally expressed as inactive precursor proteins, which require specific DUB proteases to expose their C-terminal glycine residue essential for forming an isopeptide linkage with the target protein lysine. In these cases, these DUBs (or assimilated (as some of these DUBs are not specific of Ub, but use UbIs as substrate)) express a peptidase activity. Additionally, the removal of Ub and UbIs from their target proteins is also performed by these proteases. As such, DUBs (or assimilated) play a central role in Ub/Ubl homeostasis by regulating and maintaining a sufficient pool of Ub/Ubls and also in processing them for efficient conjugation. Based on their sequence similarity, DUBs (or assimilated) can further be subdivided into five distinct subfamilies (defined in Chapter 1): USPs, UCHs, OTUs, MJDs, and JAMM/MPN+ (Fig. 3-1) [148]. The USP, UCH, OTU and MJD families correspond to cysteine-dependent proteases, whereas the JAMM/MPN+ family groups zinc-dependent metalloenzymes. The catalytic domains amongst different families of DUBs (or assimilated) reveal a large diversity in secondary structure showing distinct structural folds. Examples of canonical folds encountered in DUBs (or assimilated) include USP domains, Ubl domains and JAMM/MPN+ domains [148]. Outside of the conserved catalytic domains, these enzymes also show a large variation in the number and type of additional domains that they contain [213]. Crystal structures of representatives of each of these five DUBs (or assimilated) families have been solved, some even in complex with Ub or Ubl derivatives [149]. Some accessory domains around the catalytic region have been structurally characterised, but this remains discreet and little is known of the role of these accessory domains on the DUB (or assimilated) activity. In various cases, DUBs (or assimilated) form part of larger protein complexes. Diversity is also present in the variety of protein-protein interactions in which these DUBs are involved, as well as in the substrate specificity displayed by these enzymes [213]. They can also

hydrolyse different kinds of chemical linkage, although, sometimes, with differing efficiencies. Members of this family of enzymes can cleave ester, peptide and isopeptide bonds to Ub [214]. In the following few pages, starting from the Cys-DUBs, the different members of the DUB (and assimilated) family will be described and their means of regulation, when known, discussed.

Cysteine-dependent DUBs rely on two or three essential amino-acid residues forming a catalytic dyad or triad, respectively, also common to classical cysteine proteases, such as papain. During catalysis that comprises two main steps, the thiol group of the catalytic cysteine is responsible for performing a nucleophilic attack on the isopeptide bond of the ubiquitinated Lys residue. This activity is facilitated by the presence of an adjacent histidine side chain, which lowers the pKa of the cysteine by deprotonation. In some cases, a third residue, usually aspartate or asparagine, aligns and polarises the catalytic histidine. Enzymes that lack a third residue display other means to polarise the catalytic histidine [148]. In this reaction mechanism, an acyl intermediate is formed by the covalent linkage between the cysteine and the carboxyl group that is generated upon cleavage. The reaction results in the release of the target protein and the formation of a covalent intermediate between the cysteine protease and the Ub moiety. The intermediate formed undergoes a hydrolysis, resulting in the deconjugation of the Ub from the enzyme and the regeneration of the enzyme [149].



**Figure 3-1 Schematic representation of the active site of cysteine-dependent proteases showing the mechanism of cleavage of an isopeptide bond.**

**A.** Schematic representation of the active site of cysteine-dependent proteases showing the mechanisms of cleavage of an isopeptide bond. The catalytic cysteine is aligned with a proton-withdrawing group to facilitate nucleophilic attack on the substrate. **B.** Schematic view of the active site of metalloproteases. A water molecule is activated to serve as a nucleophile which is able to attack the carbonyl carbon of the isopeptide bond.

Structural studies on the four cysteine-dependent DUBs subfamilies have revealed both striking similarities and differences [149]. As mentioned previously, they commonly share a papain-like catalytic triad comprising Cys-His-Asn/Asp [213].

The largest and most diverse DUBs have been classified in *the USP sub-family*. USPs, ranging from 38 kDa to 370 kDa in size, contain a catalytic domain of approximately 38

kDa. The catalytic core of USP enzymes is often decorated by various other domains, including protein interaction domains or domains that determine their subcellular localisation, where examples include USP19 and USP30 that are located in the endoplasmic reticulum and the outer membrane of mitochondria, respectively [215, 216]. Interestingly, various USP enzymes contain one or multiple Ubl domains, suggesting an auto-regulatory mechanism [217]. The various structures of USP family DUBs have indicated that the USP domain fold is highly conserved despite containing low sequence similarity. The USP domain generally comprises of three subdomains, commonly referred to as the Finger, Palm, and Thumb, although examples including CYLD (for Ubiquitin carboxyl-terminal hydrolase) (DUB implicated in the human benign tumour syndrome [218]) lack the finger sub-domain. The Palm and Thumb sub-domains harbour the catalytic centre, while the Finger sub-domain accounts for the interaction with the distal Ub [148, 219]. The molecular basis and the regulation of substrate recruitment in USP proteins have been addressed with structures bound to Ub. Comparison with apoenzymes of the USP family has indicated that the correct orientation of the catalytic residues is essential for their activity. This has been characterised for USP7, where Ub binding is required to induce conformation changes that are essential for the competent positioning the catalytic cysteine in the proximity of the catalytic histidine [220, 221]. In contrast, USP14 and USP8 contain a catalytic triad that is already properly orientated in the absence their Ub substrate. However in the case of USP14, Ub binding results in the translocation of Ub-binding surface loops enabling the C-terminus of Ub to access the active site [222]. However, inactive forms of USP, due to catalytic ability or substrate accessibility, is not a global characteristic for this family of DUBs, as CYLD, for example, is both catalytically competent and structurally substrate permissive, even in the absence of its substrate [148, 223]. These findings depict that this sub-family of DUBs displays heterogeneity in its regulatory mechanism.

**The UCH sub-family** of DUBs consists of four members: UCH-L1, UCH-L3, UCH-L5, and BAP1 (BRCA1-associated protein-1). All UCH enzymes expose a conserved catalytic domain (230 amino acid residues) and, similarly to the USP family, three conserved residues, Cys, His, and Asp. UCH37 and BAP1 contain unique C-terminal extensions of lengths 100 and 500 residues, respectively. An additional BRCA1 interaction domain is present in BAP1 [224]. In contrary to USP enzymes, which are able to cleave large protein substrates, the UCH family preferentially shows activity on smaller leaving groups, for instance by-products of proteasomal or lysosomal degradation or short C-terminal extensions of polymeric Ub precursors [225, 226]. The most apparent structural feature of UCH enzymes is the presence of a large surface loop, which, upon Ub binding, folds over the active site. As such the C-terminus of Ub needs to thread through this loop which causes significant steric constraint on the size of the ubiquitinated substrate that can be cleaved [227]. Indeed, *in vitro* studies have shown that the restricted accessibility of UCH enzymes results in negligible activity against poly-Ub chains of any linkage type [152, 225]. By contrast, UCH-L5 can act on larger substrates, such as poly-Ub chains despite containing an analogous large active site loop. UCH-L5 binds to the Rpn13 subunit of the 26S proteasome via its C-terminal extension and functions in recycling Ub chains from proteasome substrates. UCH-L5 does not cleave poly-

Ub chains ‘en bloc’, but progressively shortens them [228]. This suggests that either the binding to the proteasome induces conformational rearrangement of the active site loop or that the proteasome significantly unfolds the Ub polymers allowing their entry through the obstructing loop. Taken together, the structural and biochemical data on these DUBs reiterate the subtle regulatory and specificity principles observed in DUBs.

The size of *the OTU domain sub-family* members ranges from 25 to 135 kDa, where the OTU core domain comprises 150 to 200 residues, although the catalytic core of some members can extend to 360 residues [148, 229]. This family of enzymes is largely implicated in cell signalling processes. Similar to the USP family, OTU enzymes contain additional Ub interacting motifs, Ub-associated domains, and Ubl folds. Although the structure of the inactive OTU domain is considerably different to that of the USP domain, upon activation the catalytic residues of the OTU DUBs are well superimposed with that of the USPs [230]. Indeed, as commonly observed for DUBs, active site rearrangement is also a requirement for various OTU family members and characterised examples include OTUB1, which undergoes conformational changes upon Ub binding [231]. As shown for OTU members, the recurrently occurring catalytically resting state of cysteine-dependent DUBs was proposed to be a mechanism protecting the catalytic cysteine residue against oxidative stress. At an optimum catalytic state, the low pKa cysteine is prone for attack by a reactive oxygen species (ROS), reversibly inactivating the enzyme. This has been shown to be the case for many phosphatases, as well as redox proteins. Mechanistic and structural studies have indicated that ROS predominantly target cysteine residues located in the active site. Cysteine-dependent DUBs containing low pKa cysteine residues are thus speculated to be reversibly regulated by ROS [232]. This was shown to be most likely the case of members of the OTU sub-family, where examples include the Cezanne and A20 DUBs [233, 234].

DUBs including Ataxin-3 and a number of Ataxin-3-like proteins, constitute *the MJD sub-family* [149]. Although low sequence similarity is observed between the catalytic domains of Ataxin-3 and other DUBs, structural NMR studies have indicated a conserved overall arrangement of the catalytic triad [235]. A distinct structural feature of MJD enzymes is a large helical arm that, in the absence of Ub, restricts access to the active site [235, 236]. The structure of Ataxin-3 bound to two Ub molecules further indicated that the helical arm hindering the active site regulates substrate binding. Furthermore, the back of the helical arm locates the binding site for the second Ub molecule. It has been suggested that Ataxin-3 interaction with a polymer of two distal Ub molecules is required for stabilising an open catalytic conformation [236].

*The JAMM/MPN<sup>+</sup> metalloprotease sub-family* contain a Zn<sup>2+</sup>-bound polarised water molecule. Structurally, as well as mechanistically, this family differs to the cysteine-dependent sub-families of enzymes. Metalloproteases contain a conserved, active-site comprising one aspartate, one serine and two histidine residues which together stabilise the metal through their coordination [169, 237]. The mechanism of these metalloisopeptidases is driven by the activation of the water molecule, through its binding to the zinc ion, to form a

hydroxide ion which in turn is able to attack the carbonyl carbon of the isopeptide linkage between Ub/Ubls and their target protein (which can be another Ub molecule). This reaction generates a transient non-covalent tetrahedral intermediate with the substrate. Proton transfer from a water molecule, resulting in the replacement of the amine with a hydroxyl group, is further responsible for collapsing the intermediate and releasing from the DUB. In this catalytic mechanism, mainly based on the AMSH-LP/di-Lys63-Ub crystal structure, an adjacently located glutamate residue contributes as a proton acceptor and donor (Fig. 3-1)[149, 186].

In human, eight DUBs (or assimilated) of the JAMM/MPN<sup>+</sup> sub-family have been identified, including AMSH-LP, implicated in the ESCRT machinery and in membrane receptor trafficking. Structural studies on AMSH-LP alone and bound to a Lys63-linked di-Ub allowed the analytical investigation of the JAMM/MPN<sup>+</sup> motif of metalloisopeptidases. The crystal structure of AMSH-LP (residues 264-436 (AMSH-LP numbering)) identified the JAMM motif and two insertions, referred to as Ins-1 (residues 314-339 (AMSH-LP numbering)) and Ins-2 (residues 393-415 (AMSH-LP numbering)) [186]. These structures further highlighted the AMSH-LP His347, His349 and Asp360 residues and a water molecule forming a hydrogen bond with Glu292, as the key players in Zn<sup>2+</sup> coordination. A comparable coordination system had been identified in the structure of the prokaryote *Archaeoglobus fulgidus* (Af) JAMM [237]. Additionally, AMSH-LP binds two zinc ions; one of which is implicated in catalysis, as just described, and a second one bound to the large Ins-2 loop, shown to be responsible for recognising the proximal Ub of the K63 di-Ub. These observations have grouped both AfJAMM and AMSH-LP, as part of the subset of proteins containing a JAMM/MPN<sup>+</sup> domain, which include a highly conserved EX<sub>n</sub>H(S/T)HPX<sub>7</sub>SX<sub>2</sub>D motif describing a distinct arrangement of two histidines and one aspartate preceded by a conserved glutamate [159]. Other JAMM/MPN<sup>+</sup> family proteins include the BRCC36 [168], Rpn11 [159], MYSM1 [238], eIF3f [179], and CSN5 [70], enzymes reported to have activity on Ub and/or Ubls.

Prior to structural studies, the role of the JAMM motif present in the Rpn11 subunit of the 19S proteasome lid was probed by site-directed mutagenesis of the two conserved histidines [159], confirming their requirement for catalysis. A similar approach was also used to characterise the JAMM motif of CSN5, where mutations in its conserved JAMM residues abolished or severely comprised its catalytic activity. Dependence of the CSN deneddylase activity upon the presence of divalent cations was further confirmed by metal chelators [70]. Together these results emphasised that a functional JAMM motif of metalloisopeptidases is essential for their hydrolytic activity on substrate. In general, some JAMM/MPN<sup>+</sup> also share the propensity to associate in large protein complexes, such as the 19S proteasome lid, the CSN complex, and the BRCC36-containing complexes (BRCA1-RAP80 and BRISC). Interestingly, most of these JAMM/MPN<sup>+</sup> domain-containing enzymes have been shown to lack DUB (or assimilated) activity when not incorporated within their complexes [47, 70, 168]. These observations raise the question about the regulation of DUB activities in general and a closer look into the potential mechanisms underlying their regulations is important.



Because of the importance of understanding the basis of CSN5 regulation, our group elucidated the structure of the catalytic N-terminal region of the human CSN5 subunit [63]. In addition to confirming that CSN5 is a metalloisopeptidase, the structure of this subunit displays, as often observed in various other DUBs (or assimilated), an auto-inhibited state. Low DUB activity is often observed in *in vitro* studies, suggesting the occurrence of imposed activation mechanisms in cells and possibly of activatory role of accessory domains or subunits in the case of multi-protein complexes. Additionally, in some instances, where hindrance of the substrate-binding site or where the catalytic site harbours an inactive conformation, conformational changes are necessary to occur in order to express the DUB activity. As this Chapter focuses on our work on the activity regulation of CSN5, three major regulatory mechanisms of DUBs reported in the literature are briefly presented below, illustrating strategies adopted by DUBs (or assimilated) to control their activity.

(i) Post-translational modifications. A strong relationship has been described between the phosphorylation and the Ub signalling pathways. For example, upon phosphorylation, CYLD deubiquitination activity on TNF receptor-associated factor 2 (TRAF2) is inhibited, resulting in the suppression of the NF- $\kappa$ B signalling network [239]. Various studies have suggested the possibility that Ub or Ubl modifications could regulate the catalytic activity of various enzymes, both inside and outside the DUB family. For the ATXN3 and USP25 DUBs, ubiquitination activates ATXN3 catalytic activity, whilst SUMOylation of USP25 results in steric hindrance inhibiting its activity [240, 241]. Outside of the DUB family, there are many examples and one of them is the CRLs which show enhanced activity, through their modification by Nedd8 [242, 243].

(ii) Subcellular localisation. At the cellular level, in a large amount of cases, there is an apparent coupling between subcellular localisation and enzyme activation, where the subcellular localisation can determine the substrates available for processing. This has been evident for AMSH, which shows both nuclear and cytosolic localisation, however exhibits activity solely after being recruited by its activator STAM into endosomes present in the cytosol [244]. Various other DUBs including USP8, undergo epidermal growth factor (EGF)-dependent translocation into endosomes, where its activation triggered upon associates with STAM [245, 246]. From the DUB family, only two, USP19 and USP30, encompass a transmembrane domain, whilst several show nuclear accumulation including USP36, which localises to the nucleolus and plays a role in regulating its structure and function [216, 247].

(iii) Auto-inhibition. In various cases, the activity and specificity of DUB activity is regulated via the presence of additional Ub-binding domains or motifs. Additionally, activation of many DUBs can be induced by their association with other partners, potentially increasing substrate binding. This is the case for AMSH and USP8, which are activated via their interaction with the UIM (for Ub interacting motif)-containing signal transducing adaptor molecule 2 (STAM2) making these enzymes more efficient in capturing their substrates [244]. This principle is largely used in the regulation of many other examples of DUBs, resulting in, either their activation or inhibition. For UCH37, its activity is reduced when associated with the INO80 chromatin-remodelling complex [248]. The cysteine isopeptidase USP5 contains a ZnF-UBP domain which is able to interact with a free Ub. This

results in conformational changes leading to an increase in the catalytic rate of Ub chain processing by USP5 [249]. For various metalloisopeptidases, the incorporation into large multi-protein complexes is required for their activation. This has been demonstrated in the case of Rpn11, CSN5, and BRCC36, which are part of the 26S proteasome, the CSN, and the BRCC36-containing complexes, respectively [70, 159, 160, 168]. In these examples, incorporation into large complexes has been suggested to serve as scaffolds, presenting DUBs (or assimilated) to their physiological substrates or, perhaps to play a role in regulating catalytic activity, potentially through conformational adjustments producing a catalytically competent enzyme. However, this is not the case for AMSH and AMSH-LP which do not require to associate within a complex to expose activity *in vitro*. However, in the case of AMSH, its interaction with STAM through its N-terminal domain contributes to its activity by holding Ubs simultaneously with AMSH [250].

Despite the insight obtained regarding the catalytic activity of the CSN through mutagenesis, the molecular basis for its activation and activity has just started to be revealed. The most studied and known activity of the CSN complex is its deneddylase activity on neddylated Cullins and, more particularly, on Cull1. Our main focus has been to characterise the regulatory mechanisms implicated in the activity of the CSN complex. To address this, we took an approach that encompasses a detailed investigation of the catalytic activity of this complex, mostly from the perspective of the catalytic subunit CSN5. As with other JAMM/MPN+ containing multi-protein complexes, the CSN catalytic activity rests on its metalloprotease subunit CSN5.

In mammals, CSN5, originally identified as an interactor and a co-activator of the AP-1 transcription factor, has additional interactions with factors regulating signal transduction, proliferation and survival of cells, suggesting that CSN5 may be implicated in functions reaching further than the regulation of the CRLs [129, 251]. Notably, CSN5 has been repeatedly found to interact with and to control the nuclear export of the CDK inhibitor p27 [86], the Rad9-Rad1-Hus1 complex [252], Runx3 [91], Smad7 [90], and the tumour suppressor p53 [253]. CSN5 contributes to the CRM1 (for Chromosome region maintenance 1 protein homolog; also known as Exportin-1 or XPO1)-dependent translocation of some of these binding partners from the nucleus to the cytoplasm and to their subsequent increased proteasome-dependent-degradation. The detailed understanding of the exact role(s) of CSN5 in these processes remains however to be investigated. Moreover, CSN5 that could act as a multi-functional protein appears to exist in various contexts, including, CSN5 in a stand-alone form (for more on that topic, please refer to Chapter 3 section 3.5), in smaller CSN sub-complexes, and within the CSN complex. CSN5 was originally found to exist independently of the holo-Cop9 complex in *At* [254] and was later characterised in smaller CSN subcomplexes in mammalian cells [126]. Although the deneddylase function of CSN complex has been ascribed to its CSN5 subunit, the molecular composition, activity or function of these CSN sub-complexes and of CSN5 stand-alone form, have not been reported. Importantly and as detailed in Chapter 1, substantial evidence has accumulated that demonstrate a strong link between CSN5 and cancer. This has become evident with the high

expression levels of CSN5 found in many human cancers which, in some cases, correlates to poor prognosis and to reduced expression levels of p27 [82, 251]. The implication of CSN5 as an important player in such diverse cellular processes, including phosphorylation, deneddylation, and translocation, and the fact that it is found outside of the CSN complex could potentially explain why its activity is kept under tight regulation.

For these reasons, obtaining a better understanding of the catalytic and regulatory mechanism of CSN5, within and outside the context of the CSN complex is essential. The elucidation of the CSN5 crystal structure by the group allowed, upon my arrival, to structurally and functionally investigate the catalytic subunit of the CSN complex. CSN5 has been reported in two catalytic forms, one incorporated within the CSN complex, catalytically active and the other found in a CSN-independent state, inactive. As previously described the activity regulation of DUBs (or assimilated) is essential for many members, where a recurrent principle is their auto-inhibition. To begin understanding how CSN5 expresses its isopeptidase activity, thorough structural comparison with relevant model systems, but mostly with AMSH-LP was performed. Structural analysis was complemented with biochemical, biophysical, and *in silico* characterisation. A constitutively catalytically active variant of CSN5 was designed and validated, unravelling a potential auto-inhibitory mechanistic requirement for the activity of CSN5 at the molecular level. Insights gained on the regulation of CSN5 in the stand-alone state could provide some clues as to how CSN5 harbours a catalytically competent conformation within the CSN complex.

### 3.2 Structure of CSN5:

The *csn5* gene is located on the chromosome 8q13 and is highly conserved in *Hs*, *Mm*, *Dm*, *Sp*, and *At*. Orthologous proteins of CSN5 from *Hs* and *At* share more than 60% sequence identity. The main human CSN5 protein isoform consists of 334 amino acid residues (38 kDa). Two structural motifs are present in CSN5, a nuclear export signal (NES)-like sequence [255, 256], and the catalytic JAMM motif. The NES-like sequence mediates the CRM1-dependent transportation of proteins from the nucleus to the cytosol, which has been demonstrated through CSN5 interaction with p27, while the MPN domain represents the catalytic domain of CSN5. Over the years, CSN5 has been implicated in numerous signalling pathways including those that regulate light signalling in plants, larval development in *Dm*, and integrin signalling, cell cycle control, and steroid hormone signalling in a variety of systems, although this implication might be due to its pleiotropic function of CRL regulator. Various studies have suggested that CSN5 is essential for development and survival of animal models and it was shown that *csn5* gene K.O. in mice is embryonically lethal [110].

Within the CSN complex, a second subunit, namely CSN6, harbours a MPN domain, as it is the case for other MPN domain-containing large multi-protein complexes, such as the Rpn8 subunit of the proteasome 19S regulatory lid and the eIF3h subunit of the translation initiation factor eIF3. This general organisation is illustrated in Fig. 1-5. Each of these multiprotein



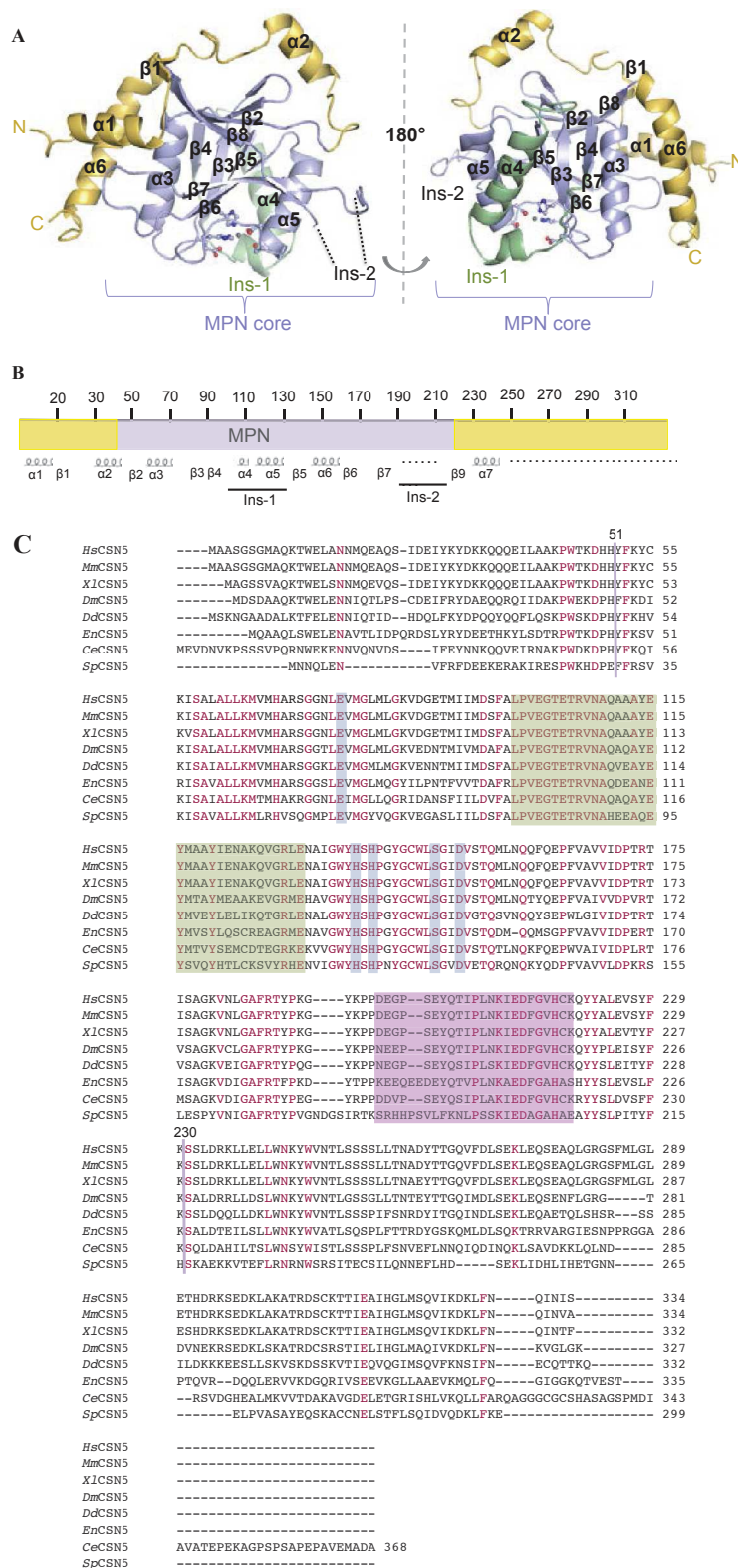
complexes displays two MPN domain-containing proteins, one belonging to the MPN+ subfamily and containing a JAMM motif (members include; CSN5, Rpn11, eIF3f) and the other from the MPN- subfamily and which lack the catalytic JAMM motif (members include; CSN6, Rpn8, eIF3h). It is noteworthy that other JAMM/MPN+ domain-containing proteins exist without an associated MPN- counterpart - this is the case of AMSH and AMSH-LP, which neither associate with MPN- domain-containing proteins, nor form multi-protein complexes with PCI domain-containing proteins. Similarly, an MPN- domain-containing protein, Prp8 lacks association with proteins of the JAMM/MPN+ family and exist within the U5 small nuclear ribonucleoprotein particle, which itself forms part of the spliceosome network [257].

### **3.2.1 CSN5**

In this part, the structure of CSN5 will be presented and key features are discussed in the light of the structure and biochemistry of other MPN family members. In the first section, the CSN5 structure containing the residues 1-257 will be considered, whilst, in the second one, the C-terminal region of CSN5 will be described through recent structural data [61]. We further aim to gain an in-depth understanding of the dynamics of the CSN5 MPN domain and identified key players in the regulation of CSN5 activity.

#### **3.2.1.1 CSN5 1-257 crystal structure**

The crystal structure of the human CSN5 isopeptidase domain (residues 1-257) comprises a core JAMM/MPN+ motif and two insertions first defined in AMSH-LP [186], Ins-1 (residues 97-131 (CSN5 numbering)), located in the vicinity of the active site and Ins-2 (residues 197-219 (CSN5 numbering)), disordered in the crystal structure (Fig. 3-2 A). The heterogeneity of the sample, as well as its low expression levels guided us to work on the fragment 1-257 which corresponds to the MPN core, flanked by an additional 50 residues extending from its N-terminus and 26 residues from its C-terminus. The structure of the 1-257 fragment of human CSN5 obtained thus lacks 77 amino acid residues at the C-terminal end. The catalytic core of CSN5 comprising the residues 51 to 230, includes an eight-stranded  $\beta$ -sheet, numbered  $\beta$ 1 to  $\beta$ 8, further completed by three  $\alpha$ -helices (Fig. 3-2 B). The topology of the JAMM motif has previously been defined by structural studies of other MPN+ domain-containing structures, including AMSH-LP. The conserved JAMM motif of CSN5 tetrahedrally coordinates a  $\text{Zn}^{2+}$  via its active site residues, H138, H140 (located on the  $\beta$ -sheet  $\beta$ 5 and on the loop immediately following it, respectively), D151 (located on the  $\alpha$ -helix  $\alpha$ 5), and a water molecule which hydrogen bonds to residues E76 (located on the loop following the  $\alpha$ -helix  $\alpha$ 3) and S148 (on the loop between the  $\beta$ -sheet  $\beta$ 4 and the  $\alpha$ -helix  $\alpha$ 5). This results in a zinc coordination site bringing together residues from the helix  $\alpha$ 5, the central  $\beta$ -sheets,  $\beta$ 5,  $\beta$ 6, and  $\beta$ 7 and additional loops including those located between the  $\beta$ -sheet  $\beta$ 4 and the  $\alpha$ -helix  $\alpha$ 5, and the Ins-1 segment (Fig. 3-2 A). The elucidated crystal structure contains additional N- and C-terminal appendices, which pack tightly against its core catalytic domain, profoundly modifying the properties of its surface (Fig. 3-2 A).

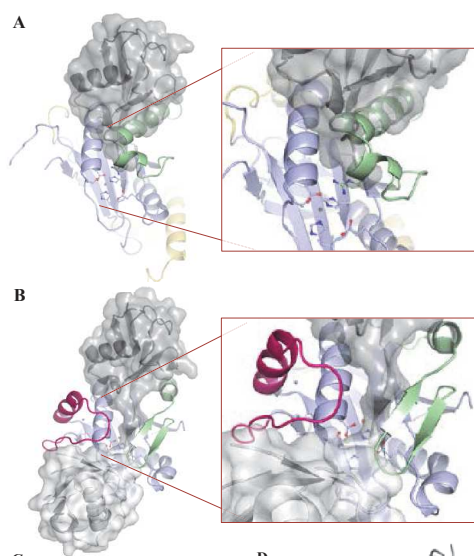


**Figure 3-2. Structure of CSN5 MPN domain.**

**A.** Overall monomeric crystal structure of CSN5 1-257 (PDB code: 4F7O) revealing its core MPN domain (in blue) decorated with peripheral N- and C-terminal extensions (in yellow). Secondary structure elements are numbered and the catalytic residues forming the JAMM motif are displayed in ball-and-stick representation. The flexible Ins-1 segment, shown in green, comprises residues 97-131 and the disordered Ins-2 region comprises residues 197-219. Right panel: 180° rotated view of left Panel. **B.** Schematic representation of the CSN5 domain organisation. The MPN domain is delineated in violet; the N- and C-terminal extensions are yellow. The Ins-1 and Ins-2 insertions are placed with respect to their location in the sequence. The secondary structure elements are indicated underneath. Regions that are either not ordered in the crystal structure or not included in the crystallised fragment are indicated by a black dotted line. **C.** Sequence alignment of CSN5 sequences from 8 diverse organisms. The sequence alignment was carried out by retrieving CSN5 sequences with Blastp using human CSN5 as the search sequence and 8 diverse sequences from various organisms were selected. The alignment of these sequences was performed by Clustal Omega server (default parameters). The Uniprot website ([www.uniprot.org](http://www.uniprot.org)) was used to carry out this alignment. *Hs*: (Uniprot entry: Q7L5N1); *Mm* (Uniprot entry: Q3UIT2); *Xl*: *Xenopus laevis* (Uniprot entry: Q6NUC2); *Dm* (Uniprot entry: Q9VCY3); *Sm*: *Schistosoma mansoni* (Uniprot entry: G4V7G0); *Os*: *Oryza sativa* (Uniprot entry: Q6ZKM2); *Bm*: *Bombyx mori* (Uniprot entry: Q2F614). The strictly conserved residues amongst the eight selected sequences are indicated in red. Highlighted in blue are the

As previously observed for the structure of AMSH-LP, the two additional features comprising the Ins-1 and Ins-2 segments also decorate the CSN5 MPN core. Sequence conservation across CSN5 orthologs indicates that the Ins-1 region is highly conserved, whilst the Ins-2 one contains a small number of conserved residues (Fig. 3-2 C). In contrast to the Ins-1 segment of AMSH-LP, which consists of a long helical portion and a pair of anti-parallel  $\beta$ -strands, CSN5 Ins-1 is mainly helical ( $\alpha 4$  and  $\alpha 5$ ) [186]. Such insertions are also present in other members containing a MPN- domain, but the Ins-2 segment in MPN- members is generally much reduced in length, compared to its counterpart in MPN+ proteins. These include CSN6, Rpn8 and Prp8, where the Ins-1 insertion of the first two, similarly to AMSH-LP, adopts a long helical segment and an anti-parallel  $\beta$ -hairpin, whilst in the latter reveals an extended conformation. The large variety of conformations adopted by the Ins-1 portions among CSN5 paralogs, supported by a large variation in their sequence could be functionally important. In AMSH-LP and AMSH, the Ins-1 segment is involved in its interaction with distal Ub [188]

The CSN5 Ins-1 segment is completely folded back over its active site (Fig. 3-2 A). As suggested by the comparison between CSN5 and AMSH-LP structures, the orientation of the Ins-1 region could hinder the access for the C-terminus of the substrate, thereby contributing to the auto-inhibitory state of CSN5 in the isolated state (Fig. 3-3). This would be consistent with the observation that CSN5 in the stand-alone form is void of catalytic activity. In our work, AMSH-LP has been used as a model system allowing the study of CSN5. The reason for this is two-fold; the structures of both the apo- and substrate bound forms of human AMSH-LP are available and, additionally, at the start of the work, it was the only MPN+ structure available that was supplemented with robust biochemical characterisation. Most recently, two studies on the structure of the *Sc* proteasome lid enzymatic subunit Rpn11 in complex with Rpn8 allowed further comparison and understanding [158] [160]. However, the Rpn11 structure in the context of a heterodimer with the Rpn8 subunit is somewhat different to that of stand-alone CSN5 and we preferred discussing these recent structural contributions in Chapter 4.



**Figure 3-3. Comparison of AMSH-LP and CSN5 Ins-1 region.**

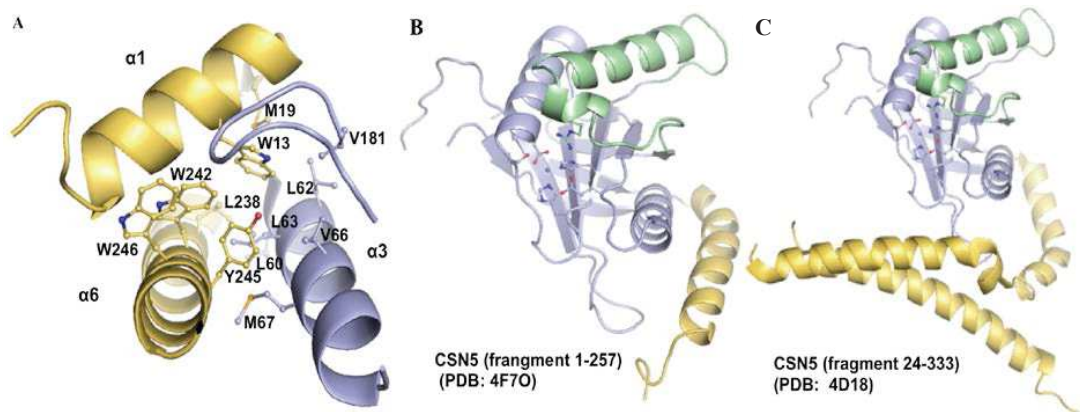
**A.** The Ins-1 insertion of CSN5 (in green) hinders the access of Nedd8 C-terminus into the active site. A model of CSN5-Nedd8 (in grey) interaction by using the AMSH-LP/Ub (PDB code: 2ZNR) crystal structure to superimpose. **B.** The crystal structure of AMSH-LP/Ub (PDB code: 2ZNR) showing the binding mode of Ub (in grey) to AMSH-LP. The Ins-2 segment is shown in magenta.

In the structure of CSN5, electron density is lacking for the region corresponding to its Ins-2 segment, indicating a highly flexible segment. A disordered Ins-2 segment is also observed in the structure of Rpn11 within the Rpn11-Rpn8 dimer [48, 158, 160]. In this work, the role of the Ins-2 region was proposed to be linked to the contact with another subunit of the 19S regulatory particle of the proteasome. The large variation in the size of the Ins-2 region amongst the various MPN family members is observed. CSN5 contains the largest Ins-2, made of 30 residues, whilst a 2-amino acid Ins-2 element is present in MPND [67]. Low levels of sequence conservation among MPN paralogs have been observed for the Ins-2 segment. In addition to differences in length, Ins-2 segments also show variation in composition and conformation, making it an interesting component of the MPN core to study. Furthermore, although well described for AMSH-LP, the role of Ins-2 in other members of the family remains largely uninvestigated.

Amongst the structures of MPN domain-containing proteins elucidated to date, a number of them have shown the propensity to form oligomeric arrangements. These include the structures of *Af* JAMM, *Hs* Rpn8 from the 19S proteasome lid, *Hs* CSN5, *Dm* and *Hs* CSN6 from the CSN complex. Despite their inclination to dimerise, these proteins assemble mostly via different surfaces [67]. In the analysis of all possible contacts within the human CSN5 crystal structure, two different types of dimerisation interfaces can be observed, organised in dimer of dimers. One of these dimers related by a local twofold symmetry axis is contained within the asymmetric unit (A-B dimer), whereas the other is generated by crystallographic operators (A-A' dimer). Structural analysis using the PISA Server [210] determined a solvation free energy for interface formation ( $\Delta G^{\text{int}}$ ) at  $-14.0 \text{ kcal mol}^{-1}$  for the A-B type dimer and at  $-10.9 \text{ kcal mol}^{-1}$  for the type A-A' dimer. The small difference in  $\Delta G^{\text{int}}$  determined (less than  $10 \text{ kcal mol}^{-1}$ ) between the two CSN5 dimers challenges the discrimination of the crystallographic artefact and serves impossible, thus further criteria had to be considered (See Section 4-5).

### 3.2.1.2 Model and structure of CSN5 C-terminal extension

Up to very recently, the studies performed on the MPN members were, in general, limited to the core MPN domains of these proteins, sometimes a little extended or reduced [63, 64, 176]. A recurrent reason for this is the problems encountered in the recombinant expression of full-length (FL) proteins, where, in most cases, their C-terminus is truncated prior to crystallisation. Evidence obtained from the lack of ordered density and from the observed truncations during expression and purification suggests an exposed highly flexible region allocated to the C-terminus of these proteins. Briefly, electron microscopy structures of the CSN complex [46] and of the 19S proteasome lid [175, 258] proposed that their MPN+ and MPN- subunits interact via their MPN domains, supported by earlier biochemical evidences [168, 259]. More recent work showed that their C-terminal regions are mainly involved in governing interaction with the other subunits within their complexes [259]. This hypothesis



was further confirmed by biochemical characterisation. Furthermore, Förster and coworkers [258] were able to position the crystal structure of Rpn8 and a molecular model of Rpn11 into their corresponding location in an electron microscopy map. This allowed to resolve a *hitherto* unassigned density consisting of long twisted helices forming a coiled-coil bundle, which extrudes away from the MPN domain of the two proteins and form contacts with other PCI subunits [258]. Moreover the role of the C-terminus of MPN domain-containing proteins within the proteasome lid was also investigated, where findings demonstrated that the C-terminus of both Rpn11 and Rpn8 are sufficient and essential for the assembly of the complex, forming a helical bundle together with the C-terminus of PCI subunits. It was also suggested that this role of the C-terminus of MPN domain containing proteins, described for the proteasome lid, may be a common feature used by the other Zomes complexes, the CSN and eIF3 complexes [59].

**Figure 3-4. N- and C-terminal extensions of CSN5.**

**A.** The N- and C-terminal extensions (in yellow) that flank the MPN core adopt a constrained topology that is stabilised by the presence of a hydrophobic core contributed by the helices  $\alpha 1$ ,  $\alpha 3$  and  $\alpha 6$ . Specific residues involved in this core are shown in ball-and-stick mode. Elements from the MPN core are shown in blue. **B.** C-terminal extensions (in yellow) for the MPN CSN5. The MPN core is represented in blue, with the Ins-1 segment in green. **C.** C-terminal extensions (in yellow) for the FL CSN5 subunit in the context of the CSN crystal structure. The MPN core is represented in blue, with the Ins-1 segment in green.



In our work, FL CSN5 can be obtained, however various degradation products are also present in the sample. In general most MPN protein structures solved till very recently lacked the C-terminal extension, with the exception of Prp8 which contains a similar length of extension to that of CSN5, however adopts a very different position and conformation to that observed in CSN5. In the structure of CSN5 that we elucidated, the C-terminal helix ( $\alpha_6$ , fragment 232 to 257) together with the N-terminal helix  $\alpha_1$  wrap around and form extensive contacts with the MPN core (Fig. 3-2 A; B; Fig. 3-4 A,B). For context, the structure of CSN5 containing the residues 24 to 333 part of the crystal structure of the CSN complex recently determined is presented in Fig. 3-4 C. This highlighted the helical content of the C-terminal portion of the protein.

### **3.2.2 A structural comparison between AMSH-LP, AMSH, and Rpn11**

In this section, we aim to present the MPN family from a structural point of view. The recent years have seen a wealth of MPN-domain structures being solved and provides a solid ground to better understand the architecture of this family and to identify structural elements potentially involved in the regulation of their functions.

#### **3.2.2.1 MPN domain**

To date, in eukaryotes, the number of proteins containing an MPN domain extends to fourteen, namely, Abraxas, Abro1, AMSH, AMSH-LP, BRCC36, CSN5, CSN6, eIF3f, eIF3h, Rpn8, MPND, MYSM1, Rpn11, Prp8, from which seven of them are JAMM/MPN+ motif containing DUBs [149]. Additionally, a number of prokaryotic MPN domain containing proteins have also been identified, these include; the phage tail assembly protein K coming from bacteriophage lambda, and its homologs found in other phages and prophages [169].

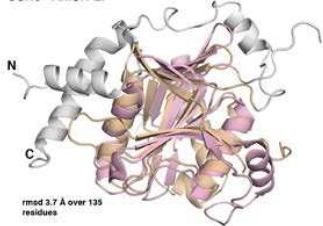

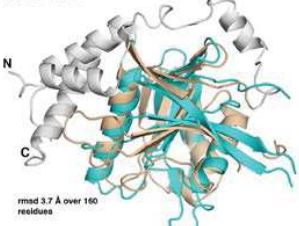

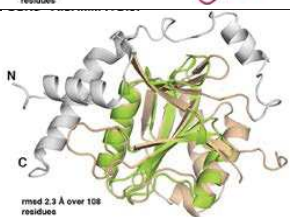
Analysis of the MPN domain-containing structures elucidated to date allow to define the MPN fold as one consisting of up to eight-stranded  $\beta$ -sheet which is further decorated by up to three  $\alpha$ -helices. Typically, the core  $\beta$ -sheets are arranged to adopt a partially outspread barrel shape, protected by helices. The MPN core was initially defined by structural works on the prokaryote AjJAMM, and has later been confirmed through the structures of Rpn8, AMSH, AMSH-LP, Prp8, CSN6, CSN5 and Rpn11. Although considerable elaboration on the AjJAMM topology is observed in eukaryote member, comparison between these eukaryote structures shows little variation in the secondary structure content between their MPN core domains. In some members, such as Rpn8 and Prp8, the  $\beta_7$  and  $\beta_8$  strands are fused together, resulting in MPN domain containing structures which consist of seven  $\beta$ -strands.

In our work, structural comparison of the MPN domain present in the JAMM/MPN+ family of proteins, AMSH-LP, AMSH, Rpn11 and CSN5, was used to understand this family of

proteins (Fig. 3-5 A-C). The structures of CSN5 and other MPN domain-containing proteins share a core, as highlighted in Table 3-1 that present the root mean square deviation (rmsd) values of pairwise structural superimposition carried out with Chimera.

**Table 3-1. Rmsd values across the MPN sub-family against CSN5 (PDB code: 4F7O).**

Calculated in Chimera. In the four columns, CSN5 MPN core and the N-/C-terminal extensions are shown in beige and grey, respectively. Prp8 MPN appendices are also displayed in grey.

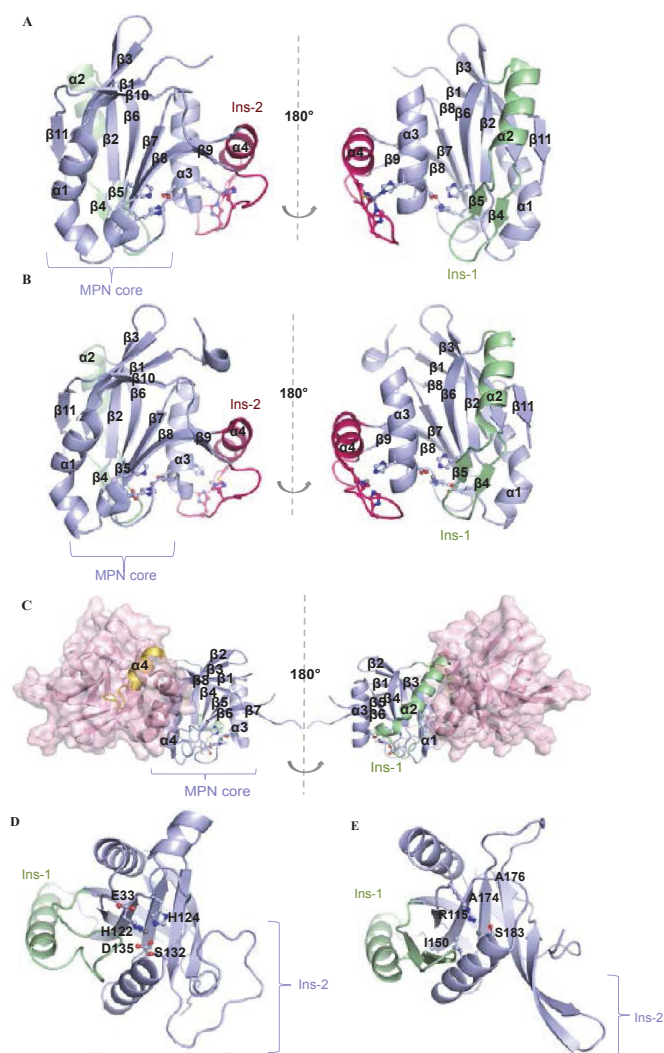
CSN5/	Rmsd value (Å)	Number of C $\alpha$	Superimposition
AMSH-LP	3.7	135	
<i>Dm</i> CSN6	2.2	64	
<i>Hs</i> Rpn8	3.7	160	
Prp8	2.9	137	
<i>Af</i> JAMM	2.3	108	

The major differences present in the structures of these proteins locate on their Ins-1 and Ins-2 segments. These elements show diversity in both length and composition.



### ***3.2.2.2 Molecular models and analyses of other MPN+/JAMM proteins***

To date from the seven JAMM/MPN+ DUBs known, only AMSH-LP, AMSH, and more recently CSN5 and Rpn11 (in complex with Rpn8) have solved crystal structures of their MPN domains. Here we built molecular models, using Phyre2 [260] of the unknown structures including that of BRCC36 and eIF3f and carried out structural analysis to compare JAMM/MPN+ proteins of known and unknown structures (Fig. 3-5. D,E). Although most of the MPN domain can be highly superimpose with other JAMM/MPN+ proteins, major differences are observed in the Ins-1 insertion of both modelled proteins, compared to the Ins-1 regions of proteins such as CSN5 and AMSH-LP that have been crystallised. The differences observed in the Ins-1 region are both structural, as well as showing varying topology. The variation in the Ins-1 region both in proteins have been solved and also in the models, presented here suggests that the Ins-1 is a flexible region that can adopt various conformations. We further analyse the flexibility of these insertions in the following sections.



**Figure 3-5. The JAMM/MPN+ members.**

The crystal structures or models of JAMM/MPN+ are presented; the MPN core is shown in blue, the Ins-1 region in green and the Zn-binding site residues are represented in ball-and-stick. Right panel: 180° rotated view of left Panel. **A.** AMSH-LP; PDB code: 2ZNR **B.** AMSH; PDB code: 3RZU **C.** Rpn11 (in complex with Rpn8); PDB code: 4O8X **D.** and **E.** Models of BRCC36 and eIF3f MPN+ proteins built using Phyre2 [260].

### 3.2.2.3 Structural elements: Zinc-binding site, Ins-1, Ins-2

The JAMM motif observed in the DUBs of the JAMM/MPN+ family and which is found in all three domains of life (bacteria, archaea, eukarya), has suggested that this motif specifies a catalytic centre defining this family of metalloisopeptidases [149]. This has now been validated on a number of JAMM/MPN+ members (Fig. 3-5). Sequence analysis allows to identify a number of conserved polar residues within the MPN domain of the MPN+ subset. The initial studies determining the characteristic architecture of this motif explored the structure of A/JAMM [65, 66]. The JAMM motif typically contains a catalytic  $\text{Zn}^{2+}$  which is coordinated via a tetrahedral manner through the contribution of the Ne2 of two histidine residues located on the strand of one the main  $\beta$ -sheets, the carboxylate from an aspartate residue situated on the  $\alpha 2$  helix, and a water molecule (Fig. 3-1). Additionally, a glutamate residue located on the loop between  $\alpha 1$  helix and  $\beta 2$  strand serves as an acid-base catalyst. A highly conserved serine residue present between the histidine ligands is able to form a

hydrogen bond interaction with the main chain of the glutamate and stabilises the intermediate state. Found in two positions upstream from this serine residue accommodates an aromatic residue, commonly a tryptophan. Overall these residues form a pattern of EX<sub>n</sub>H(S/T)HPX<sub>7</sub>SX<sub>2</sub>D specific for the MPN<sup>+</sup> subset of protein.

Similarly to CSN5, the structures of AMSH-LP, AMSH, and Rpn11, all display, in their stand-alone forms or Rpn8-bound form in the case of Rpn11, the conserved JAMM motif (Fig. 3-6). Being the sole JAMM/MPN<sup>+</sup> isopeptidases to have been structurally elucidated in both its apo- and its substrate bound states, AMSH-LP and, more recently, AMSH, provide a model system to understand and compare the catalytic competence and enzyme-substrate interactions of metallo-dependent isopeptidases [186, 261]. Structural comparison of the zinc-binding sites of these proteins reveals that the global topology of their active site is maintained (Fig. 3-5; 3-6). This observation can be further complemented by analysis of the distances between the Zn and the catalytic residues of these isopeptidases (Table 3-2). MD simulations were performed on CSN5 and AMSH-LP, where the dynamics of the side-chain of the zinc coordination implicated residues was analysed. For the simulations, here and thereafter, the dummy atom approach was used to mimic the tetrahedron-shaped zinc divalent cation. As such four peripheral dummy atoms are attached to the central zinc atom. This approach is expected to impose orientational constraints to the four zinc ligands (H138, H140, D151, and water). Using this approach, the position of the catalytic residues over the course of the 50 ns simulations was analysed, as a control ensuring that the zinc environment was maintained. The obtained values were coherent with those from the crystal structure (Table 3-2; Fig. 3-6). Moreover, the stability of the zinc-coordinating site was analysed by choosing non-coordinating residues in the vicinity of the site, S148 and W146 for CSN5 and S357 and F355 for AMSH-LP. We observed that both pairs of residues selected are maintained at a steady distance from the Zn throughout the simulation (Fig. 3-6 E and F). These results highlight the conserved topology of their Zn-binding site and the neighbouring area. Interestingly, however, they do not all show catalytic activity, whilst AMSH-LP is active, CSN5 in the stand-alone form is void of catalytic activity (data shown in section 3.4). Comparison with the constitutively active AMSH-LP MPN domain dictates that the observed topology of the Zn-binding site in CSN5 and Rpn11 is likely to be in the optimal position and geometry for pursuing isopeptidase activity. This in itself indicates that the JAMM motif of CSN5 in isolation is catalytically competent and further suggests that other features contained in these proteins are implicated in regulating their activity. A striking difference observed for this region between CSN5 and AMSH-LP (and AMSH) structures is the presence, in the former, of a salt bridge formed between the R106 residue, belonging to its Ins-1 segment, and the Zn-coordinating D151 residue. Sequence conservation analysis of CSN5 orthologs revealed a highly conserved Ins-1 and a strictly conserved R106 residue. This type of electrostatic interaction is lacking in other MPN-domain containing proteins, suggesting a potential CSN5 unique regulatory mechanism.

Exploring the topology of the JAMM motif in other MPN enzymes, the model of BRCC36 displays a canonical JAMM motif (Fig. 3-5 D). However, for the model produced for eIF3f, the JAMM motif appears highly divergent (Fig. 3-5 E). With the exception of the catalytic serine residue, the remaining JAMM motif residues, including the two histidines are replaced by an alanine and a glycine respectively, and the glutamate and the aspartate are replaced by an arginine and an isoleucine. Findings indicating that eIF3f lacks the canonical Zn-coordinating residues had previously suggested that this subunit and thus the eIF3 complex lack DUB activity. More recently, however, it was reported that the eIF3 has an intrinsic DUB activity on activated Notch receptor [179]. Further structural and biochemical evidence is required to establish the DUB activity of the eIF3 complex and to understand the lack of JAMM motif residues.

The zinc-binding site is a unique and conserved feature of the MPN+ members of the JAMM/MPN+ family. These proteins contain another two important features, as briefly previously presented, the Ins-1 and Ins-2, which are probably responsible for the important difference in regulation and in activity displayed by these enzymes. Although commonly present in MPN domain-containing proteins, these two structural elements exhibit a large variability showing highly divergent insertion sequences, structural composition, as well as potentially functional properties.

**Table 3-2. Distances between Zn and ligand atoms in the active site of the CSN5 JAMM motif.**

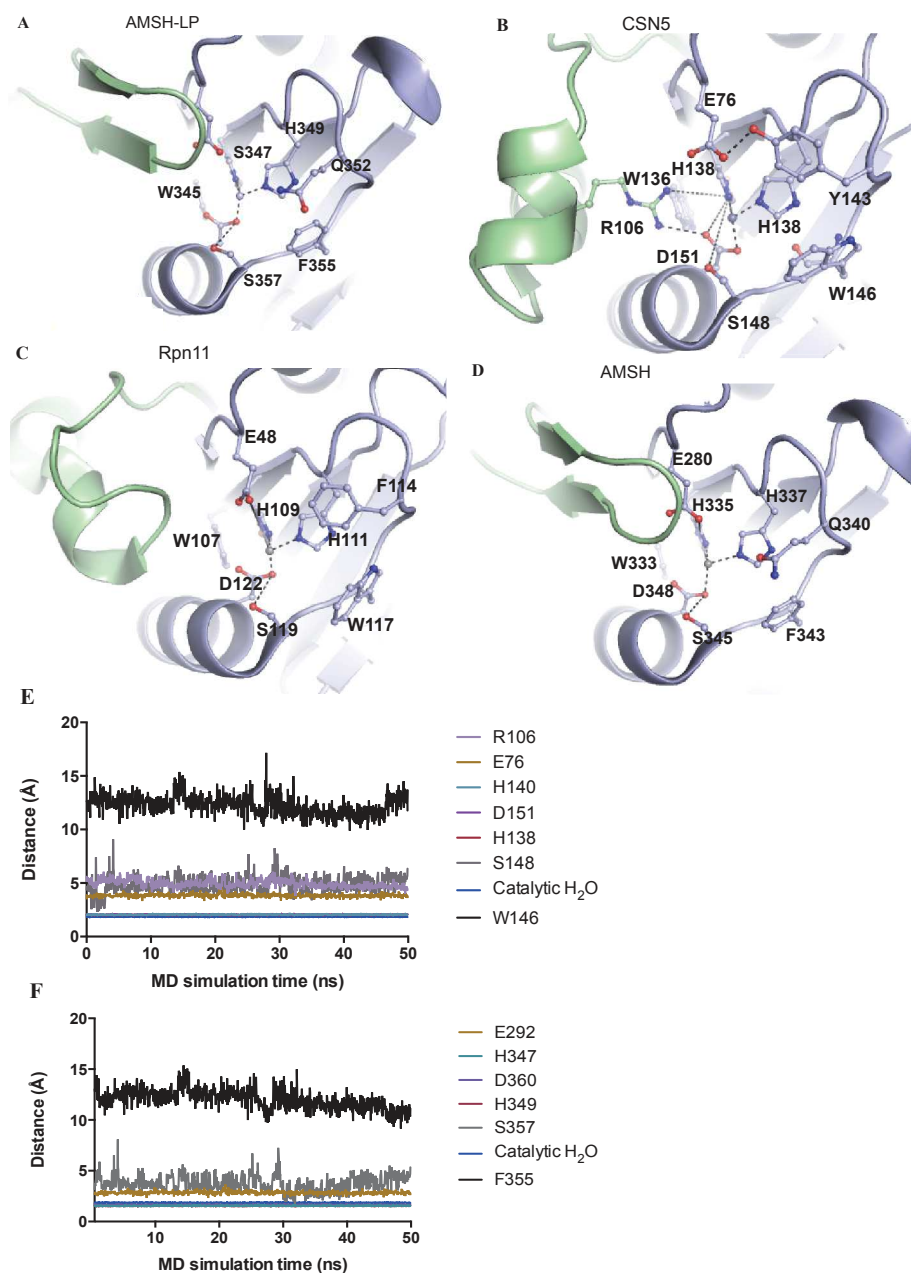
The comparison between CSN5 1-257 crystal structure, and the Zn-dependant DUB AMSH-LP (PDB code: 2ZNR; [186]) and their corresponding MD simulation are presented. The first four ligands belong to the tetrahedral coordination sphere. Residue numbers are indicated in bold for CSN5.

Atom distance (Å)	CSN5 1-257		AMSH-LP	
	X-ray model (A/B)	MD	X-ray model (A)	MD
Zn NE2 - <b>H138</b> /H347	2.07 / 2.10	2.03 ± 0.03	2.04	2.01 ± 0.02
Zn NE2 - <b>H140</b> /H349	2.05 / 2.03	2.05 ± 0.03	2.02	1.98 ± 0.03
Zn O - <b>Wat</b>	2.13 / 2.15	1.86 ± 0.02	1.92	1.63 ± 0.02
Zn OD2 - <b>D151</b> /D360	2.06 / 2.02	2.02 ± 0.03	1.94	1.90 ± 0.03
Zn CD - <b>E76</b> /E292	5.35 / 5.31	5.03 ± 0.25	4.71	4.50 ± 0.03
Zn OG - <b>S148</b> /S357	3.49 / 3.71	3.87 ± 0.30	4.32	4.40 ± 0.02
Zn CZ - <b>R106</b>	5.15 / 4.83	5.13 ± 0.20	-	-
Zn NE1 - <b>W146</b>	12.45 / 11.75	11.73 ± 0.20	-	-
Zn CZ - <b>F355</b>	-	-	12.3	11.8 ± 0.06

In both AMSH and AMSH-LP, their Ins-1 insertion forms a short two-stranded anti-parallel  $\beta$ -hairpin and contributes to the catalytic site of the enzyme. Although structurally very different to the Ins-1 of AMSH-LP/AMSH, the Ins-1 regions of AMSH-LP/AMSH and of CSN5 are all in the immediate vicinity of the substrate binding site. Structurally, the Ins-1 of CSN5 is composed of a helical segment,  $\alpha 4$ , and a loop forming between  $\beta 4$  and  $\alpha 4$ . Superimposition with the Ins-1 segment of AMSH-LP, AMSH, and Rpn11 with the Ins-1 of

CSN5 reveals major differences in topology and occupancy within the active site. Most interestingly, CSN5 contains a crucial arginine residue (R106) located on its Ins-1 segment and which further projects to establish a salt bridge with the catalytic Asp151 residue. This feature of CSN5 has not previously been observed in other JAMM/MPN+ family members. Thus further characterisation of this salt bridge was required to understand its function and for questioning its presence in CSN5.

Only the Ins-2 regions of AMSH-LP and AMSH have been structurally defined. The Ins-2 region of AMSH-LP is composed of an  $\alpha$ -helical element followed by an extended loop harbouring the residues C402, H408 and H410, which, together with H362 (extending from the JAMM core), are implicated in the coordination of a second, non catalytic  $\text{Zn}^{2+}$ . The substrate-bound crystal structure of AMSH-LP allowed to ascribe a function of its Ins-2 region that was validated by enzymatic probing. It is implicated in recognising the proximal Ub of di-Ub moieties. Although not applicable for JAMM/MPN DUBs that do not bind di-Ub moieties, AMSH-LP Ins-2 contributes to catalysis, possibly by positioning correctly the substrate into the active site (mutations in the Ins-2 of AMSH-LP principally affect the  $k_{\text{cat}}$  of the hydrolytic reaction) and by participating to the linkage specificity of AMSH-LP [186]. A similar conformation of Ins-2 is also observed in Prp8, although, in this case, it does not coordinate a second  $\text{Zn}^{2+}$ . In CSN5 and Rpn11, these regions are disordered in the structures obtained. The reason for the lack of a definable conformation may be a result of the absence of a secondary  $\text{Zn}^{2+}$  coordinating site, which, in contrast, is present in the structures of both AMSH and AMSH-LP. The  $\text{Zn}^{2+}$  binding residues in these enzymes are implicated in the structuring of their Ins-2 loop, which in turn allows its interaction with the proximal Ub. For the Ins-2 insertion of Rpn11, a distinct role has been defined and is distinct to that of AMSH-LP and AMSH's. The unresolved Ins-2 segments in both the Rpn11 and CSN5 structures may additionally suggest that this region contains high flexibility and that its conformation may depend on its interaction with neighbouring partners within their complexes. This was shown, very recently, to be indeed the case for CSN5 in the context of the CSN complex [61].



**Figure 3-6. The catalytic MPN core domain.**

Close-up view of the JAMM motif of AMSH-LP (**A**), AMSH (**B**), Rpn11 (**C**) and CSN5 (**D**). Residues in the vicinity are shown in stick and ball representation. The Ins-I region is indicated in green. **E**. Fluctuations of CSN5 zinc binding site during MD simulations. Time dependence of the distance between the catalytic zinc, the side chains of the active site (H138, H140, D151, S148, R106, E76, W146) and water molecule during the MD simulations. **F**. Fluctuations of AMSH-LP zinc binding site during MD simulations. Time dependence of the distance between the catalytic zinc, the side chains of the active site (H347, H349, D360, S357, E292, F355) and water molecule during the MD simulations.



### **3.2.3 Molecular basis of CSN5 auto-inhibition**

#### ***3.2.3.1 Implication of AMSH-LP and of AMSH Ins-1 in substrate recruitment***

Auto-inhibition has been an ascribed characteristic for various members of the JAMM/MPN<sup>+</sup> family, however AMSH-LP and AMSH are key examples of isopeptidase enzymes that exist in a competent conformation even when unbound to their di-Lys63-Ub substrate. Examples of inactive MPN<sup>+</sup> proteins include Rpn11, BRCC36 and CSN5. These three enzymes have been shown to require integration into their respective host complex(es) to express their catalytic activity. As described earlier, comparison of the conserved active site implicated residues between AMSH-LP with CSN5 or Rpn11 reveals that the catalytic residues align very well and contain a common overall topology. This suggests that, in the case of CSN5 and Rpn11, most probably, other factors are implicated in determining and forming a catalytically competent enzyme. Variation in the core MPN domain of these proteins has been highlighted and located mostly to their two insertions, Ins-1 and Ins-2, although additional variations beyond the construct borders used in structure determination could not be excluded, reflecting an important caveat in the structural biology of protein domains. As a result of its location in the active site of JAMM/MPN<sup>+</sup> proteins, the Ins-1 region can be considered as a key participant. Across members of the family, this region reveals large diversity in composition, in their secondary structure composition and conformation and in their function.

In AMSH-LP, the Ins-1 region composed of an antiparallel  $\beta$ -sheet and a long  $\alpha$ -helix contributes to an extensive part of the substrate binding groove (Fig. 3-3). The structure of AMSH-LP bound to its di-Ub substrate allowed to define the function of this insertion in being implicated in interacting with the distal moiety of its substrate.

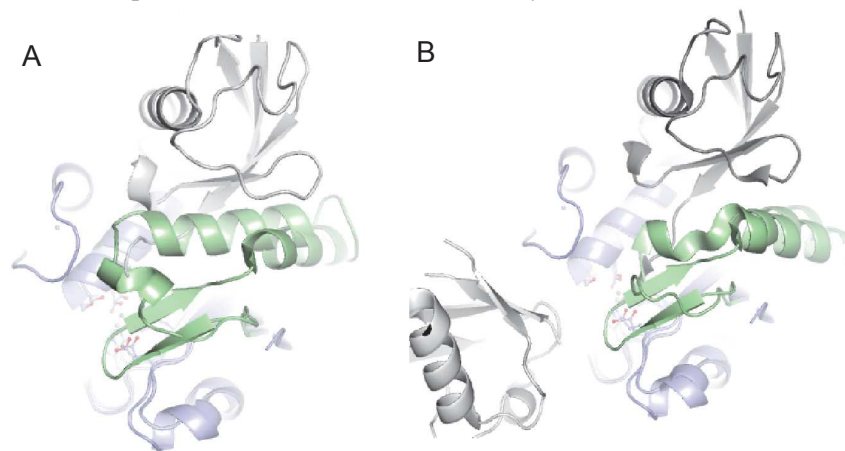
#### ***3.2.3.2 Auto-inhibition of CSN5 through the Ins-1 conformation***

Here we report the investigation of the active site of CSN5 that allowed to gain insight in the regulation of its activation. CSN5 has been found in two different states of activity, one where it is integrated in the CSN complex and displays deneddylase active and the other one where it is not integrated in the CSN complex and is catalytically inactive. The conserved overall topology of MPN<sup>+</sup> active site suggests that CSN5 and AMSH-LP bind their corresponding substrate (Nedd8 and Ub, respectively) in a similar orientation. This hypothesis is supported by the high residue conservation observed in the catalytic region between the two proteins and is further suggested by the high sequence identity (60%) between Nedd8 and Ub.

To identify determinants that might explain the lack of activity of CSN5 in the stand-alone form, we assembled a crude model of CSN5 bound to Nedd8 in two steps. First, CSN5 monomer structure (PDB code: 4F7O) was superimposed onto the structure of the DUB



domain of AMSH-LP/di-Lys63-Ub (PDB code: 2ZNV). Second, the structure of Nedd8 (PDB code: 1NDD) was superimposed onto that of the distal Ub moiety from the AMSH-LP/di-Lys63-Ub structure. The assembled CSN5-Nedd8 complex represented in Fig. 3-7A revealed that, in the conformation of the stand-alone form, CSN5 is unable to accommodate its substrate. Interestingly, the Ins-1 region of CSN5 containing a predominant helical character sits over the potential Nedd8-binding groove. The distance between side chains of residues located in the inner surface, including R106, V107, and T103, of the Ins-1 segment and the catalytic Zn atom identified values ranging from 4 Å for R106 to 9 Å for T103 and 11 Å for V107. As a result the topology and position of the Ins-1 insertion covering over the catalytic groove hinders the access for the Nedd8 C-terminus into the active site, highlighting for the first time a possible reason for CSN5 inactivity in the stand-alone form.



**Figure 3-7. CSN5 conformation in the stand-alone form is not compatible with Ubl binding.**

**A.** Superimposition of CSN5 structure (PDB code: 4F7O) with the structure of AMSH-LP (PDB code: 2ZNR). The conformation of their Ins-1 (in green) is compared. **B.** Superimposition of Rpn11 structure (PDB code: 4O8X, in complex with Rpn8) with structure of AMSH-LP (PDB code: 2ZNR). The conformation of their Ins-1 (in green) is compared.

The produced model of the complex CSN5 1-257/Nedd8 clearly depicts clashes between the C-terminus of Nedd8, residue 71-76, and regions of the Ins-1 segments. The tip comprising the residues 104-106 of the Ins-1 region and its following  $\alpha$ -helical element contribute to the observed hindrance. Solely based on this structural observation, one can suggest that, for CSN5 to be able to accommodate its substrate and thus be activated, its Ins-1 region requires substantial structural rearrangements. Given the different orientation of the long Ins-1 helix between AMSH-LP and CSN5, these conformational changes are likely not to be only local but probably engage profound distance rearrangements of this region. Moreover, the  $\alpha$ -helix located at the C-terminus of its Ins-1 segment would require flexibility in order to be able to relocate to a configuration more reminiscent to that of the Ins-1 of active AMSH-LP and therefore compatible with Nedd8 binding (Fig. 3-3; 3-7A).

Comparison between the Ins-1 region of Rpn11 with that of AMSH-LP and CSN5 suggests that the orientation of this region in Rpn11 is more closely related to the one observed in AMSH-LP (Fig. 3-7B; 4-8A). In a recent development in which the structure of *Sc* Rpn11, in complex with Rpn8 was determined, it was also shown that Rpn8-bound Rpn11 displays hydrolytic activity [158, 160]. Unexpectedly, Rpn11 Ins-1 region in that context adopts a rather different overall conformation that would be more closely described as a mixture between AMSH-LP and CSN5's Ins-1 conformations. However, the segment 78 to 81 of Rpn11 Ins-1 is also sterically precluding the catalytic groove for substrate binding. On the basis of these observations, Martin and co-workers investigated the effect of deleting residues of the Ins-1 loop, abolishing hindrance of the active site. Their work showed that the deletion results in the complete inhibition of Rpn11 DUB activity and they further hypothesised that, in fact, this region is of great importance for Rpn11/Rpn8 activity [160]. This study proposed that these residues are potentially important for catalysis and perhaps substrate positioning. However, further experimental proof would be required to claim that, in the conformation observed in crystal structures, this region is not acting in an inhibitory manner. More specifically, their  $\alpha$ -helical components superimpose to a better degree and occupy the same area, when compared to this element in CSN5 (Fig. 3-7; 4-8 A). Furthermore, the Ins-1 loop, specifically the side chains of residues V77 and S79 facing the vicinity of the zinc binding site of Rpn11 sit in average 6 Å away from the Zn, which when compared to the 4.3 Å distance observed between R106 and the Zn in CSN5 can explain the slightly more relaxed conformation observed in Rpn11 in the context of Rpn8. This is further shown by the lower extent of steric hindrance observed with our model of Ub docked into the active site of Rpn11 (Fig. 3-7B). Whilst we see that the CSN5 residues 104 to 119 hinder the interaction with the C-terminus of Nedd8 and that the C-terminal helix of the Ins-1 requires to translocate in order to accommodate Nedd8, for Rpn11 this helix is in a compatible conformation and clashes are observed between the C-terminal Nedd8 tails and residues 79 to 81 of the Ins-1 region.

### 3.2.4 Dynamics of CSN5

In our work, the flexibility of the Ins-1 region of CSN5 was further exacerbated by focusing on a unique feature of this region in this enzyme, which includes the presence of the salt bridge between the Ins-1 residue R106 and the Zn-coordinating residue D151 (Fig. 3-6B). Sequence conservation amongst CSN5 orthologs identifies that this active site salt bridge is highly conserved. This may indicate a functional role of this feature corresponding to the activity regulation of CSN5. The strong anchorage of the Ins-1 insertion to the Zn-binding site is suggestive of a required activation step for which one of the switches is likely to correspond to the R106-D151 salt bridge, which would allow the Ins-1 insertion to undergo restructuring to form a substrate-binding groove in which the substrate would insert. To probe this hypothesis, in the next section we investigate the flexibility of the MPN domain

and the dynamic nature of this insertion and try to characterise its role in the activity regulation of CSN5.

Two MPN structural elements, namely the Ins-1 and Ins-2 regions appear as particularly flexible. We put forward the hypothesis that CSN5 inactive to active state transition, at least partially, rests on the necessity for Ins-1 to be intrinsically malleable. In our work, detailed structural comparison and MD simulations were used to gain insight in these flexible regions. Furthermore we address the potential functional importance of these regions mainly in the context of CSN5, but which could possibly apply to other JAMM/MPN+ family members. MD simulations were performed on CSN5 and AMSH-LP to allow comparison between a catalytically inactive and a catalytically active MPN domain containing protein. The presence unique to CSN5 of a salt bridge across the substrate binding site was investigated and its implication in the flexibility and function of the Ins-1 region was revealed. Our work allowed to determine an active site lock which could be essential for the activity regulation of the auto-inhibited CSN5.

#### **3.2.4.1 Regions with high flexibility signs**

In general, the wealth of conformations that, both the Ins-1 and Ins-2 regions, adopts across the MPN sub-family have been described in the various structures elucidated to date; thus suggesting the highly flexible nature of these insertions. In various studies, the Ins-1 region has been characterised both structurally and functionally in numerous members of MPN domain-containing proteins (Fig. 3-8A and B; [63, 158, 160, 186]).

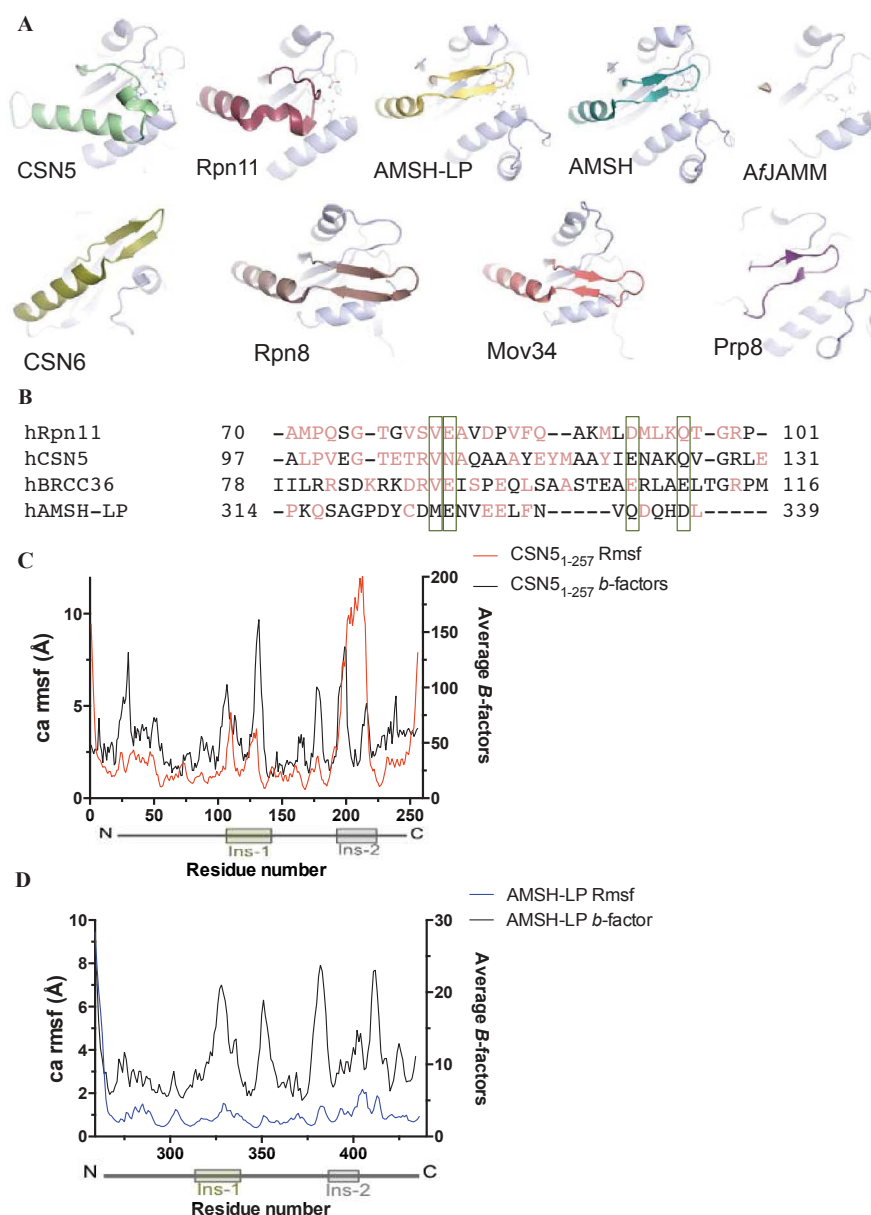
For the core MPN domain, an overall rigid structure, where in average residues show a difference in fluctuation of 1 to 2 Å (Fig. 3-8C and D), when compared to the initial crystal structure, can be depicted for both CSN5 and AMSH-LP. In general, the distribution of crystallographic *B*-factor values obtained for MPN domain-containing proteins would indicate a lower flexibility of their core domains compared to exposed regions. To identify the most flexible regions of both CSN5 and AMSH-LP, the root-mean-square fluctuation (rmsf) of the C $\alpha$  atoms were calculated by fitting the backbone atoms of all snapshots obtained during the 50 ns long MD simulations. This analysis allowed us to determine the intrinsic flexibility of the backbone of each protein. The rmsf value of each residue and the crystallographic *B*-factors are displayed on Figure 3-8C. At a first look, CSN5 in general contains a more dynamic structure compared to that of AMSH-LP (Fig. 3-8C and D). For AMSH-LP, no distinct regions of major flexibility can be depicted. The large fluctuations observed in the N-terminal end, observed frequently in MD simulations, are due to the unstructured component. In CSN5, the most flexible regions (i.e. rmsf above 2 Å) comprise residues at its N- and C-terminal regions and also residues within its Ins-1 insertion and the complete Ins-2 insertion. The high crystallographic *B*-factor values, as well as the heterogeneity observed in the different conformations that the Ins-1 region adopts in the different MPN domain-containing proteins is coherent with this (Fig. 3-8A, C and D). Interestingly, the N- and C-terminals of CSN5 are helical elements, which in the crystal structure are stacked over each other, and the

interface contains a tryptophan-rich region (Fig. 3-2A; 3-4A). Analyses from the simulations also indicate that these two helices are maintained in the same proximity to each other and, as described later in Chapter 4, could have some consequences on CSN5 regulation.

The Ins-2 region, in the crystal structure of CSN5, is disordered, which could ultimately suggest a flexible segment. For the simulations, an initial molecular model of the Ins-2 region constituted of 24 residues, was built with Modeller 9v program (<http://salilab.org/modeller>). At the time where this work was performed, structural insight into the Ins-2 region of MPN+ proteins, specifically for those not implicated in the coordination of a second Zn, was largely lacking. The high flexibility observed by rmsf calculations may be a result of the conformation of this freely floating loop at the beginning of the simulation (Fig. 3-8C). At the end of the 50 ns long simulation runs, no tendency could be ascribed to the motion of the Ins-2 region, however this potential event could not be completely ruled out for a longer simulation runs. Indeed, to observe such a drastic conformational rearrangement of this segment would require to overcome a large energy barrier. Additionally, it could also have been possible to presume that, similarly to the case with Rpn11, that the Ins-2 region of CSN5 is also implicated in interactions within the CSN complex. Moreover, the low flexibility revealed for the Ins-2 region of AMSH-LP by our simulations, may indicate that the Ins-2 of CSN5 may actually be a less dynamics component than what is depicted in the simulations, as it may either be folded back towards the core of the protein or it may be implicated in protein-protein interactions. This is further discussed in Chapter 6 in the light of recent structural data obtained in the laboratory of Dr. N. Thöma.

Specifically for the Ins-1 region, the rmsf calculations depict flexible regions located at the N-terminal (residues 108-116) and C-terminal (residues 120-132) region of this insertion (Fig. 3-8 C). These findings are coherent with our predictions from structural analysis, from which we hypothesised that the Ins-1 region requires conformational rearrangement to accommodate Nedd8. If the Ins-1 of CSN5 is to undergo a conformational change to adopt a conformation equivalent to Ins-1 in AMSH-LP, a displacement of the C-terminal helix would be foreseen, and is reflected here with the high degree of flexibility observed in the simulations.

Interestingly for CSN5 we observe that, during the course of the simulation, the R106 residue is also remained at a constant distance around 5 Å from the Zn. This is interesting because, although rmsf analysis shows flexibility in the Ins-1 region, the R106 residue is stably anchored to the Zn-binding site during the entire simulation time. This is possibly a result of its interaction with D151 and the stable salt bridge formed between these two residues. The same helical component in AMSH-LP shows lack of such flexibility. However, this C-terminal portion of the Ins-1 insertion is likely to play a role in CSN5 activation, as it adopts radically different position/orientation in CSN5 and in AMSH-LP, as well as in the Rpn8-activated Rpn11 structure (Fig. 3-7A;B).



**Figure 3-8. CSN5 MPN domain and specific extensions.**

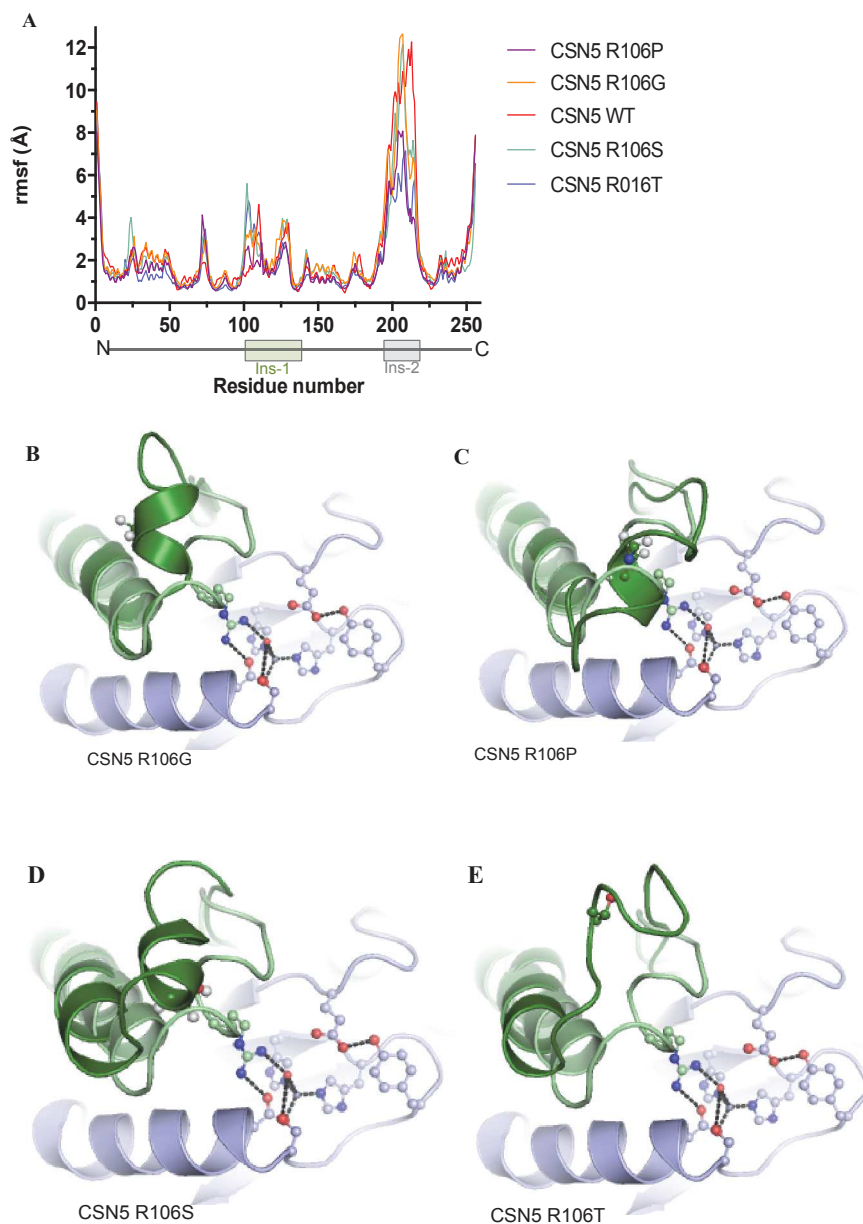
**A.** Conformational heterogeneity of the Ins-1 region amongst MPN domain containing proteins. Our comparison, includes CSN5 (PDB code: 4F7O), CSN6 (PDB code: 4EOQ), Rpn8 (PDB code: 2O95/2O96), Rpn11 (PDB code: 4O8X), AMSH-LP (PDB code: 2ZNR), AMSH (PDB code: 3RZU), Prp8p (PDB code: 2OG4), AfJAMM (PDB code: 1O10), and Rpn8 (PDB code: 2KCQ) Ins-1 segments. **B.** Sequence alignment centered on the Ins-1 region of four human MPN+/JAMM representatives, CSN5, Rpn11, BRCC36, and AMSH-LP performed using ClustalW2. Conserved residues across the species are marked in red and a green box is used to highlight conservation of residues of strongly similar properties. Overall flexibility of MPN domain containing proteins analysed by MD simulations. The average Ca rmsf obtained over the 50 ns simulations for CSN5 (**C**) and AMSH-LP (**D**) are shown.

### 3.3.3 Unlocking the active site lock of CSN5

To further question the role of the R106-D151 salt bridge in CSN5, *in silico* mutations were performed on the R106 residue to determine whether the disruption of the salt bridge triggers conformational changes of the Ins-1 segment compatible with CSN5 activation (i.e. substrate binding) and what the effect of this disruption is on the dynamics of the remaining structure. As such the arginine was substituted by different residues threonine, serine, glycine, and proline, consecutively forming various salt-bridge deficient variants of CSN5 to study. The rationale behind the use of different R106 variants comes from the observation that, in AMSH-LP, the Ins-1 is deeply involved in substrate binding and therefore sampling different physico-chemical properties at this position was of interest. A good example of the dramatic role of these residues in enzymatic activity comes from the work in [160]. The simulations of each CSN5 variant were run for 50 ns and the trajectories were analysed by calculating their rmsf plots (Fig. 3-9. A). These results showed that the core elements maintain common dynamics in comparison with the wild type CSN5. Previously, we described the dynamics of the CSN5 to be highly flexible in the regions including the Ins-1 and Ins-2 regions. Whilst this overall pattern was also observed for the CSN5 variants designed, variation in the extent of flexibility was observed. Our simulations clearly depict that, in the presence of the Ins-1 variants R106T and R106S, an increase in Ins-1 flexibility is displayed, as implied by an increase in the rmsf values of residues 99 to 104 (Fig. 3-9A). On average, the residues in these region show an increase of rmsf ranging from 1.5 to 5.0 Å difference compared to the wild-type simulation (Fig. 3-9A). Interestingly, these residues are located just before position 106. Indeed, analysis of the output trajectories from the simulations of CSN5 R106T and R106S reveals that Ins-1 of these variant shows an equilibrium between forming a 'open' and the closed form observed in the crystal structure (Fig. 3-9D and E). Further analysis of the Ins-1 region of these variants reveals that in comparison to the wild type, flexibility on residues 105 to 113 is reduced for these variants, whilst the dynamic motion of the C-terminal end of the Ins-1 region previously identified is conserved for the variants as well. This approach allowed us to identify CSN5 variants, namely R106T and R106S with the capacity to containing a higher flexibility in their Ins-1 region by disrupting the R106-D151 salt bridge. In the following section, we investigate the role of the CSN5 variants on the activity of CSN5.

In the present section we have used structural and computational approaches to reach the conclusion that the rearrangement of the Ins-1 insertion is required for CSN5 to accommodate Nedd8.





**Figure 3-9. Relaxation of the CSN5 Ins-1 conformation.**

**A.** Fluctuations of CSN5 in the wild type and Ins-1 variant forms during MD simulations. Ca rmsf fluctuations of monomeric CSN5 as a function of residue number during the 40 ns MD simulations are shown for the wild type (red), the R106T (blue), the R106G (orange), R106S (green) and R106P (violet) forms. Large fluctuations in the 190-220 correspond to the Ins-2 region that is disordered in the crystal structure and that has been modeled for the MD simulations. A snapshot after 20 ns simulations taken for each CSN5 variant, R106G (**B**), R106P (**C**), R106S (**D**), and R106T (**E**) is superimposed to the starting structure. Distinct changes in the conformation of the Ins-1 helix in variants R106T and R016S can be seen.



### 3.4 Activation of an auto-inhibited CSN5

As suggested from our *in silico* study, mutations performed on R106, part of its Ins-1 segment, that disrupt the salt-bridge trigger conformational relaxation of the Ins-1 loop. This suggested that this portion is indeed implicated in the conformational changes required for CSN5 to accommodate its substrate. This is also suggested when superimposing the simulated CSN5 R106 mutants with the active AMSH-LP crystal structure. To further validate these observations, the same mutations were performed *in vitro*, and their effect on CSN5 activity was evaluated by enzymatic measurements. In this work, the activity of CSN5 and its variants was tested against two synthetic substrates and was further verified using its physiological substrate, neddylated Cul1 in complex with Rbx1. Additionally, the binding affinity of the variants to Alexa488-labelled pro-Nedd8 was determined to allow the interpretation the activity results obtained.

#### 3.4.1 Isopeptidase activity of CSN5

With the objective being to gain functional understanding on the flexibility of the Ins-1 segment of CSN5, enzymatic activity assays were performed. Synthetic fluorogenic substrates that mimic the isopeptide bond involving the C-terminus of Nedd8 have widely been used to investigate the hydrolytic and deneddylase activity of CSN5 [262, 263]. CSN5's catalytic activity was evaluated on the two following substrates: LRGG-AMC and Nedd8-AMC. The LRGG-AMC probe represents a peptide chain corresponding to the last four C-terminal residues common to Ub and Nedd8, conjugated to AMC, whilst the Nedd8-AMC probe consists of the conjugation of the AMC with the C-terminal glycine of Nedd8. The fluorescence intensity (at the following wavelengths:  $\lambda_{\text{excitation}} = 380 \text{ nm}$ ;  $\lambda_{\text{emission}} = 460 \text{ nm}$ ) concomitant with the release of AMC, after cleavage of the amide bond between the AMC and its conjugate is proportional to the enzymatic activity. Therefore, the isopeptidase activity of CSN5 on this substrate was monitored by the measurement of the fluorescence intensity increase.

As previously been reported, the activity assays ran on CSN5 1-257 WT, using both probes, confirmed that the WT form of CSN5 lacks isopeptidase activity (Fig. 3-10 A). Moreover, as expected from our structural and computational analyses, we detected that substitution of R106 with a threonine imposes a detectable basal activity on CSN5. Using Nedd8-AMC, we also considered the other variants of CSN5, including R106 substituted with an alanine or a glycine, which also confers an activated form to CSN5, compared to the WT. However, although the activatory effect of these substitutions on CSN5 is strong, the effects R106A and R106G on CSN5 activity are more modest compared to the R106T variant. Furthermore, the R106P variant is void of isopeptidase activity (Fig. 3-10 A).

To specifically ascribe catalytic activity of CSN5, we designed a dead-mutant variant of CSN5. To do so, comparison of AMSH-LP with CSN5 allowed to identify that the glutamate 292 residue of AMSH-LP, which had been shown to lack DUB activity, aligns with the glutamate at position 76 of CSN5 [186]. The E292A variant of AMSH-LP was validated as a dead-mutant variant and used as negative control [186]. This conserved glutamate residue (E292 and E76, in AMSH-LP and CSN5, respectively) is a key component in the catalytic mechanism of JAMM motif containing DUB, playing a significant role in contributing as a proton acceptor and donor, activating the water molecule. Based on this rationale, the corresponding mutation was performed on the Glu76, substituting it with an alanine residue, for both CSN5 WT and its R106T variant. Comparison between the activity of the E76A variants with that of CSN5 WT and R106T, respectively, on the Nedd8-AMC substrate indicates a lack of activity (Fig. 3-10 A). The activity of the conformationally active R106 variants of CSN5 can also be compared to the activity observed by the CSN holo-complex, which displays a significantly higher isopeptidase activity on Nedd8-AMC substrate. An apparent 79-fold difference in the activity between CSN5 R106T and the CSN complex can be observed (Fig. 3-10 A). Further kinetics characterisation of CSN5 R106T activity (i.e.  $k_{cat}$  and  $K_M$ ) could unfortunately not be performed, because of the lack of substrate saturation over the concentration range used (in the case of CSN5 R106T).

The activity data obtained is also in agreement with the results from the MD simulations performed on the different variants of CSN5 and support our hypothesis that CSN5 in the stand-alone state is maintained in an inactive state through the closed conformation of the Ins-1 segment. The MD simulations suggested that the salt-bridge of CSN5 WT is maintained throughout the 50 ns of simulation, keeping the Ins-1 region in the proximity of the active site. This postulates a rigid conformation of the Ins-1 loop, which together with the activity data obtained, can explain the lack of activity observed by CSN5 WT. The simulations performed on the R106T variant of CSN5 demonstrated that disruption of the salt bridge lead to placing the Ins-1 insertion in a dynamic equilibrium between the opened and closed state of CSN5.

#### **3.4.2 Deneddylation activity of CSN5**

The most studied activity of the CSN complex is its ability to deneddylate neddylated Cul1 which corresponds to the removal of the Nedd8 molecule from the Cullin subunit of CRLs (presented in Chapter 1). To probe whether the enhanced activity observed with the R106T variant of CSN5 on synthetic substrates extends to its physiological substrates, its activity was tested against neddylated Cul1 complexed with Rbx1 (Nedd8-Cul1-Rbx1). Previous studies investigating the deneddylation activity of the CSN complex have shown that this complex is very effective at deneddylating Cul1 [70, 72]. This was further confirmed in our work, where we followed the deneddylase activity on Nedd8-Cul1-Rbx1 by an anti-Nedd8 Western blot detection with the decrease in neddylated Cul1 and the consecutive increase in released free Nedd8. This was carried out by gel shift assay, as described in the Material and Methods

Chapter. Interestingly the comparison of CSN5 WT and CSN5 R106T activity towards this substrate was coherent with our findings using the LRGG-AMC and Nedd8-AMC synthetic probes (described in section 3.4.1). CSN-independent CSN5 WT is unable to deneddylate neddylated Cul1, whilst the R106T variant shows a modest, however, significant, increase in the level of released Nedd8 (Fig. 3-10 B).

These results further suggest that the conformational changes occurring on the Ins-1 region of CSN5 as a result of the disruption of the R106-D151 salt bridge result in an active CSN-independent CSN5 variant. However, comparison of the activity of this variant with that observed for the CSN complex points out, as expected, that the CSN is a much better enzyme for the physiological Nedd8-Cul1-Rbx1 and the synthetic Nedd8-AMC substrate. This is not surprising and can be explained by the fact that the CSN is a complex molecular system consisting of eight subunits that are tailored for Cullin deneddylase activity. Indeed recent studies have contributed in understanding their individual roles. The EM study of the CSN complex both alone and in complex with SCF succeeded in giving insight into the topological organisation and structural arrangement of this complex [46]. Structural and biochemical analysis obtained from this work revealed that the interaction between the CSN and SCF predominantly involves the CSN subunits 1, 2, 3 and 5, although recent data also implicate CSN4 [61]. More specifically the C-terminus of the Cullin subunit was found to interact with CSN2 and CSN5 and the N-terminal conjugates of the Cullin contact CSN1 and CSN3. A general observation is that the SCF contains a broad interface which interacts with the CSN complex and which does not directly involve CSN5. Furthermore, it was shown that the presence of CSN5 is not a requisite for obtaining a CSN-SCF super-complex [46]. This implies that the interaction that CSN5 forms with Nedd8 is not dominant in the formation of the CSN-SCF super-complex. This data suggests that the contribution of Nedd8 in the affinity of SCF for the CSN is marginal, although work by the Deshaies' group suggests that Nedd8 is a mild inhibitor of CRL deneddylation [72]. The modest activity that we obtain when considering the R106T variant of CSN5 compared to that obtained by the CSN complex could stem from the weak affinity between CSN5 R106T and Nedd8, as the bulk of CSN-SCF affinity is likely to reside in interactions between other elements of these complexes. This was investigated further and is described in Section 3.3.3.1 of Chapter 3. Other factors, such as, the other CSN subunits, are therefore likely to contribute directly in the catalysis, either by increasing the affinity for the substrate, or by positioning the attached Nedd8 molecule correctly into the active site of CSN5, or both.

Our work, including the release of the R106-D151 salt bridge and inducing a conformational relaxation on the Ins-1 region, allowed to unravel one layer of the activation mechanism involved in obtaining a catalytically competent CSN5 molecule. Together these data suggest that CSN5 within the CSN complex may undergo conformational changes, activating this enzyme. Although, with the CSN5 R106T variant, we identify that the relaxation of the Ins-1 region is required for the CSN5 to be activated, it is likely not be the only requirement and

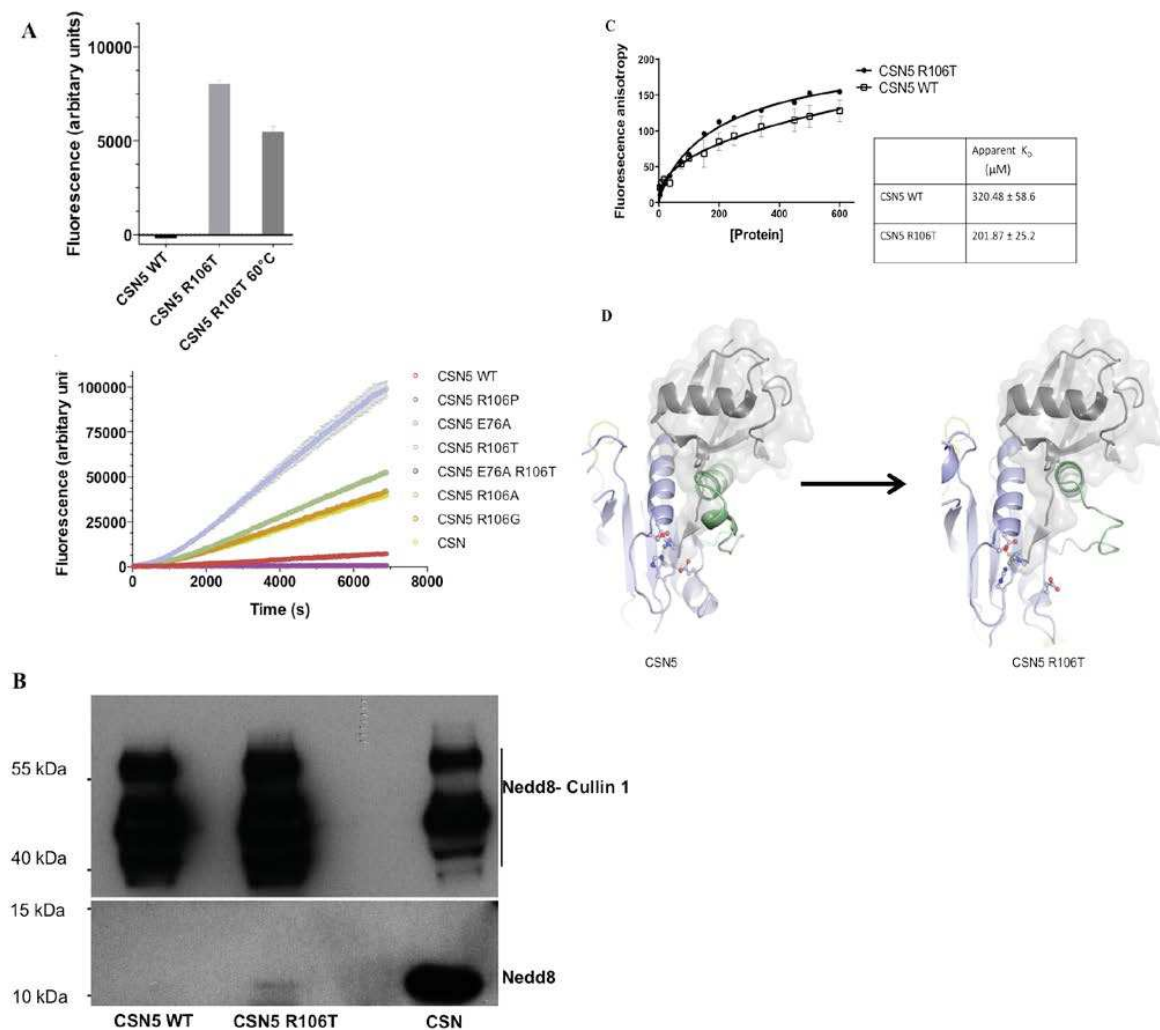
other factors may also play a role. These may include the other subunits neighbouring CSN5 within the CSN complex. Taking a closer look to the EM map of the CSN complex suggests that, for example, CSN6 is in close proximity to CSN5 [46]. Furthermore, this is coherent with data obtained from analytical biochemistry studies showing that CSN5 and CSN6 interact and with mass spectrometry, which identified the presence of CSN5 in a subcomplex containing CSN4, CSN6 and CSN7 [128, 259].

### **3.4.3 Molecular basis of CSN5 inactive state**

#### **3.4.3.1 Nedd8 binding**

Having shown that the Ins-1 region requires conformational relaxation for CSN5 activity, the underlying hypothesis to this was that the stand-alone conformation of Ins-1 prevents Nedd8 binding. To probe this idea and complement the activity data obtained, fluorescence anisotropy assays were performed to gain further understanding in the mechanism by which the R106T variant of CSN5 has an increased ability to process its substrates. The rationale behind these experiments was to determine whether the R106T variant has an increased affinity or whether it is a better enzyme for its substrate. To investigate this, the affinities of CSN5 WT and R106T, for Alexa488-labelled pro-Nedd8, were compared (Fig. 3-10C). The use of pro-Nedd8 (i.e. Nedd8 precursor that contains a 5-amino acid extension in C-terminal) is down to the fact that it is easily recombinantly produced and can be mutagenised if necessary. More work that we have performed on pro-Nedd8 can be found in Chapters 5 and 6. As suspected, CSN5 WT has a low affinity for Nedd8, displaying an apparent dissociation constant ( $K_D$ ) of around  $320 \pm 59 \mu\text{M}$ . In general, for DUBs, the values for their affinities for their mono-Ub substrates span a wide range, but are commonly characterised as weak interactions, ranging from  $1 \mu\text{M}$  to several  $100 \mu\text{M}$  [264]. More specifically for MPN domain containing proteins, Surface Plasmon Resonance (SPR) experiments showed that the pre-mRNA splicing factor Prp8, an MPN- domain-containing protein, which lacks a JAMM/MPN+ motif, has a binding affinity to Ub of around  $380 \mu\text{M}$  [265].

Interestingly, for the CSN5 R106T variant, the  $K_D$  value obtained is slightly lower than the WT form, around  $202 \pm 25 \mu\text{M}$ . Compared with CSN5 WT, these results indicate that the conformational relaxation induced on the Ins-1 segment via the disruption of the R106-Asp151 salt bridge promotes Nedd8 binding. This is observed by the increase in affinity of CSN5 R106T for Nedd8. The fluorescence anisotropy results, together with the activity data (described in Sections 3.4.1 and 3.4.2) highlight the increase in the catalytic efficiency observed with the CSN5 R106T variant. From the MD simulations performed, the flexibility of the CSN5 variants was previously determined which seems to favour the binding of Nedd8 to CSN5 (Fig. 3.10D). These results may indicate that, within the CSN complex, CSN5 undergoes structural rearrangements, one which involves the restructuring of the Ins-1 loop, either via its neighbouring interacting partners or by the substrate itself, or both.



**Figure. 3-10. Isopeptidase activity of CSN5.**

**A.** Activity of CSN5 and Ins-1 variants using the synthetic substrates LRGG-AMC (top) and Nedd8-AMC (bottom). **B.** Deneddylation activity of CSN5 and the Ins-1 variant R106T. The R106T variant form of CSN5 is able to deneddylate Nedd8- Cullin 1. Upper band shows Nedd8 signal of the Nedd8-Cullin 1 and the lower band is the Nedd8 released upon Nedd8-cullin1 isopeptide bond hydrolysis by CSN5 WT, CSN5 R106T and the CSN complex. (left to right). **C.** Fluorescence anisotropy assay showing the increased affinity of CSN5 R106T variant for Nedd8 substrate. The apparent dissociation constant ( $K_D$ ) for the active CSN5 variant was obtained at  $\pm 200 \mu$ M. **D.** Molecular models of CSN5 R106T variant recruiting Nedd8 substrate. The breakage of the R106-D151 salt bridge and relaxation of the Ins-1 is coherent with the increase in activity observed.

### 3.5 MPN homo-dimerisation

Biochemical studies, as well as crystal structure determinations of some MPN domain-containing proteins (*Hs* CSN5, *Hs* and *Dm* CSN6, *Hs* Rpn8, *Af*JAMM) have indicated their propensity to homodimerise. This is the case both for MPN+ and MPN- domain-containing proteins, where examples include CSN5 and CSN6, Rpn8, *Af*JAMM, respectively. However for others members, including AMSH, AMSH-LP and Prp8, there is no direct evidence of homodimerisation reported to date. For the members that display inclination to dimerise, these proteins assemble via different surfaces [67]. In relation to our work, the crystal structure of the human CSN5 is composed of two types of dimers, as briefly described in section 4.2.1.1 of Chapter 3.

Intriguingly, in addition to being a key component of the CSN complex, CSN5 is often found outside of this complex. Moreover, rather intriguingly, various cellular protagonists, some involved in the regulation of signal transduction and cell proliferation and survival, have been shown to interact with CSN5, as briefly described in Chapter 1 and section 3.1 of Chapter 3. This suggests that CSN5 may be involved, possibly in addition to CRL regulation, in other biochemical functions, including the specificity of transcription factors and the control of the distribution and turn-over of signalling proteins. The molecular organisation of CSN5 in its diverse functions remains unclear, however stand-alone species, as well as ones present in small CSN sub-complexes have been described [110, 126, 127]. The described ‘mini-CSN’ complexes containing a subset of CSN subunits, generally including CSN5, could hold specific physiological roles, as suggested in [127] in cell cycle control, but could also be mere stigma of CSN assembly and dissociation pathways. Additionally, these assemblies could be mediating protein-protein interactions with CSN5 partners or regulating CSN5 incorporation within the CSN complex. The fact that CSN5 organises into dimers in its crystal structure raises the question whether dimerisation of CSN5 might be observed in solution and *in vivo* and therefore potentially related to these functions.

Furthermore, in the context of large protein complexes, MPN domain-containing proteins appear to have the tendency to associate in heterodimers. The MPN domain-containing proteins, MPN+ and MPN-, have previously been identified to interact. The interaction between these two proteins is thought to be mediated via their MPN domains. Evidence of their interaction was obtained from EM data on the 19S proteasome lid [174], the CSN complex [46, 266] and the eIF3 complex [267], where their respective MPN domain-containing proteins were located in close proximity to each other.

Taken together, these elements prompted us to further probe MPN dimerisation in the context of CSN5 self-association. Advanced understanding on the basis of MPN homo-dimerisation was investigated with the rational to shed light onto the molecular basis of heterodimer formation of MPN domain-containing proteins. We extended this study in cell extract, but *in vivo* experiments were beyond our reach.



*In silico*, biochemical and other biophysical techniques including analytical ultracentrifugation (AUC; in collaboration with Dr. Christine Ebel, Institut de Biologie Structurale, Grenoble) were used to study the dimerisation of CSN5. This work led to the identification of a preponderant dimeric arrangement. Co-immunoprecipitation experiments on mammalian cell extracts performed by our collaborator, Dr. Francois-Xavier Claret (Monroe Dunaway Anderson, University of Texas, Houston) has confirmed that CSN5 homodimer is present in human cell extract and that the A-B dimer is the most likely one to exist. The presence of CSN5 dimeric or oligomeric assemblies in cellular extracts raises questions of the functional relevance of CSN5 dimeric assemblies. Moreover, the diverse interactions in which CSN5 is involved have implied that its activity could require tight regulation. It would thus be of great importance to investigate whether the dimerisation of CSN5 is part of its regulatory mechanism regarding its incorporation into the CSN complex, CSN5 dimerisation is involved in CSN-independent functions of CSN5 or the auto-association of CSN5 is related to its propensity to associate with another MPN domain containing protein, CSN6 (further discussed in Chapter 4).

### **3.5.1 CSN5 homodimerisation in crystal structure**

In the crystal, the CSN5 1-257 fragment forms two types of dimers, providing two different dimerisation interfaces; the A-B interface within the asymmetric unit and the A-A' interface between symmetrically related subunits (Fig. 3-11A). According to the PISA Server [210], the A-B and A-A' dimers display similar estimated values for free energy of dissociation  $\Delta G^{\text{diss}}$  (Table 3-3). Furthermore, analysis of the buried and accessible surface area (BSA and ASA, respectively) of the two dimers revealed also similar values (summarised in Table 3-3).

In the A-B dimer, the CSN5 monomers are related by a local twofold axis running perpendicular to the crystallography dyad axis (Fig. 3-11A). The Ins-1 segment from each monomeric species is implicated in forming part of the interface of this dimeric arrangement. (Fig. 3-11B) Additionally, the C-terminal helix (residues 232 to 241) of the opposing monomer is packing against the helical portion (residues 119 to 126) of the Ins-1. As described previously, the investigation of the flexibility of the Ins-1 insertion, within our work, highlighted that the relaxation of this segment is a pre-requisite for the recruitment of Nedd8 and to retrieve an active CSN5 subunit. With this in mind, the arrangement of the A-B dimer suggests that it could block the structural rearrangement of Ins-1 due to its contribution to the interface. The second A-A' dimer contains an interface involving monomer A residues that are implicated in the active site together with some located on neighbouring loops (residues 74, 141 to 143, 173, and 174) and the C-terminal extension, comprising the residues 248 to 255 of monomer A'. The C-terminal tail of monomer A' docks into the active site of the symmetry-related molecule A and adopts a conformation which would sterically hinder substrate binding (Fig. 3-11C). These two dimeric models A-B and A-A' thus stabilize the inactive form the CSN5 enzyme.



**Table 3-3. CSN5 interface surface properties observed in the crystal structure.**

The average ASA and the BSA of each protomer are calculated with PISA [210] for each species: dimer A-B generated by local 2-fold symmetry axis ( $\kappa=179.7$  deg, rmsd = 0.62 Å on 227 C $\alpha$  atoms), dimer A-A' generated by crystallographic dyad axis and the D2 tetramer constructed from these two dimers. The CSS interface score for dimer A-B is 0.183 and 0.136 for dimer A-A'. CSS ranges from 0 to 1 as interface relevance to complexation increases. Residues in bold play a major contribution to the interface (H-bonds, large buried surface area).

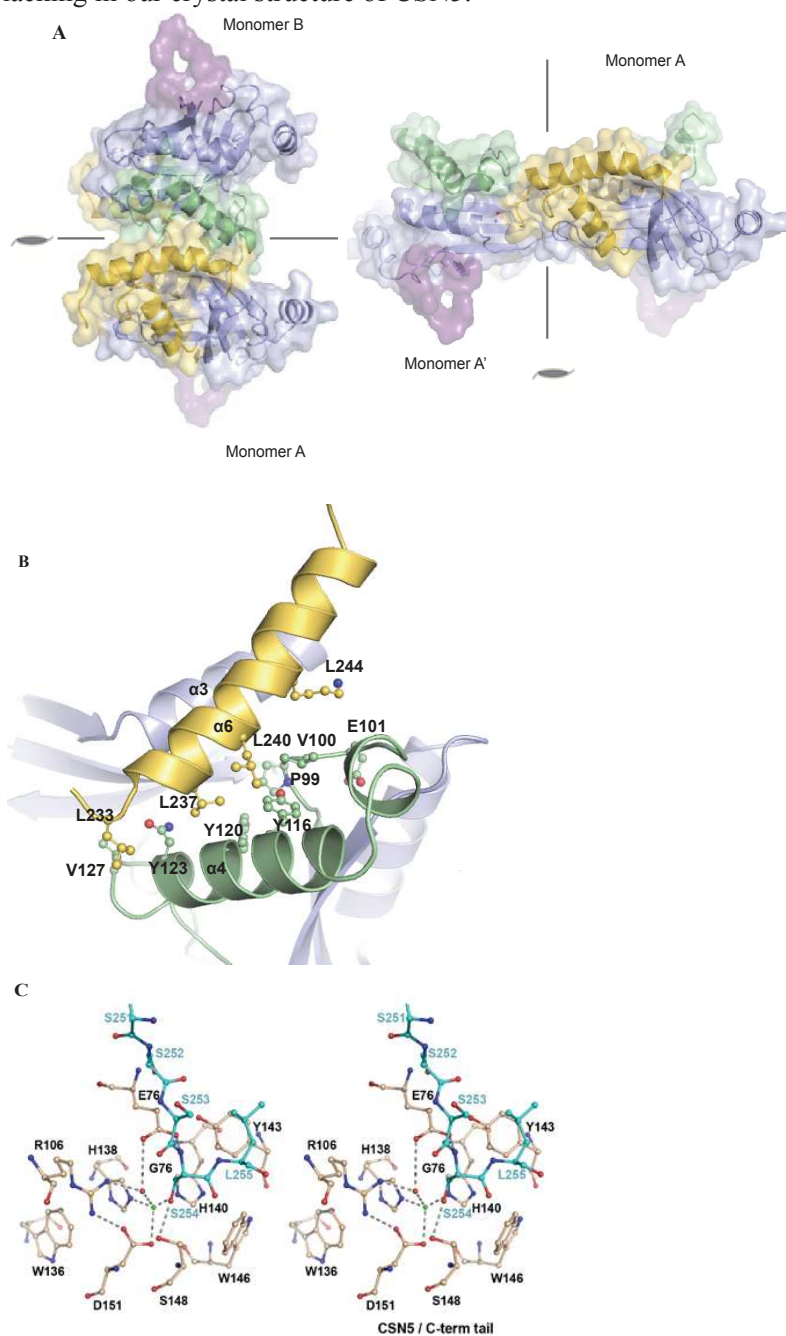
	BSA and ASA (Å <sup>2</sup> )	Number of interfacial residues	$\Delta G^{\text{int}}$ (kcal mol <sup>-1</sup> )	$\Delta G^{\text{diss}}$ (kcal mol <sup>-1</sup> )	Fraction of non-polar buried area (%)	Number of H-bond contacts	Residues implicated in the interfaces
dimer A-B	1,056 12,270	34	-14.0	3.6	59	4	L60; <b>K64</b> ; M67; L82; K84; <b>M93</b> ; D94; F96; <b>P99</b> ; <b>V100</b> ; <b>E101</b> ; E104; T105; E115; <b>Y116</b> ; M117; A118; A119; <b>Y120</b> ; E122; <b>N123</b> ; A124; Q126; <b>V127</b> ; G128; <b>R129</b> ; L130; E131; N132; <b>L233</b> ; D234; K236; <b>L237</b> ; <b>L240</b> ; L241; N243 <b>K244</b> ; V247; N248
dimer A-A'	975 12,270	32	-10.9	2.2	28	11	R70; S71; G73; <b>N74</b> ; L75; <b>E76</b> ; G102; <b>T103</b> ; R106; V107; H138; H140; P141; G142; <b>Y143</b> ; W146; S148; I150; <b>D151</b> ; D171; <b>T173</b> ; <b>R174</b> ; N248; <b>L250</b> ; <b>S251</b> ; <b>S252</b> ; <b>S253</b> ; <b>S254</b> ; <b>L255</b> ; T257
tetramer A-A'-B-B'	2,047 12,270	66	-51.3	15.8	48	15	S24; L60; <b>K64</b> ; M67; R70; S71; G73; <b>N74</b> ; L75; <b>E76</b> ; L82; K84; <b>M93</b> ; D94; F96; <b>P99</b> ; <b>V100</b> ; <b>E101</b> ; G102; <b>T103</b> ; E104; T105; R106; V107; E115; <b>Y116</b> ; M117; A118; A119; <b>Y120</b> ; E122; <b>N123</b> ; A124; Q126; <b>V127</b> ; G128; <b>R129</b> ; L130; E131; N132; H138; H140; P141; G142; <b>Y143</b> ; W146; S148; I150; <b>D151</b> ; D171; <b>T173</b> ; <b>R174</b> ; <b>L233</b> ; D234; K236; <b>L237</b> ; <b>L240</b> ; L241; N243 <b>K244</b> ; V247; N248; <b>L250</b> ; <b>S251</b> ; <b>S252</b> ; <b>S253</b> ; <b>S254</b> ; <b>L255</b> ; T257

## **Annex**

### **Paper 1**

**Insights into the regulation of the human COP9 signalosome catalytic subunit, CSN5/Jab1.**

Structural comparison between the A-A' CSN5 dimer with the di-Ub bound AMSH-LP structure may also suggest that the C-terminus positions in a similar manner to the isopeptide linkage present in the di-Lys63-Ub substrate, which is accommodated in the active site of AMSH-LP (Fig. 3-11C). Additionally, the packing of the C-terminus in the active site of CSN5 adopts a constrained horse-shoe-like conformation sterically compatible to position the remaining residues of the C-terminus, which are lacking in our crystal structure of CSN5.



### 3.5.2 CSN5 homodimerisation in solution and *in vivo*

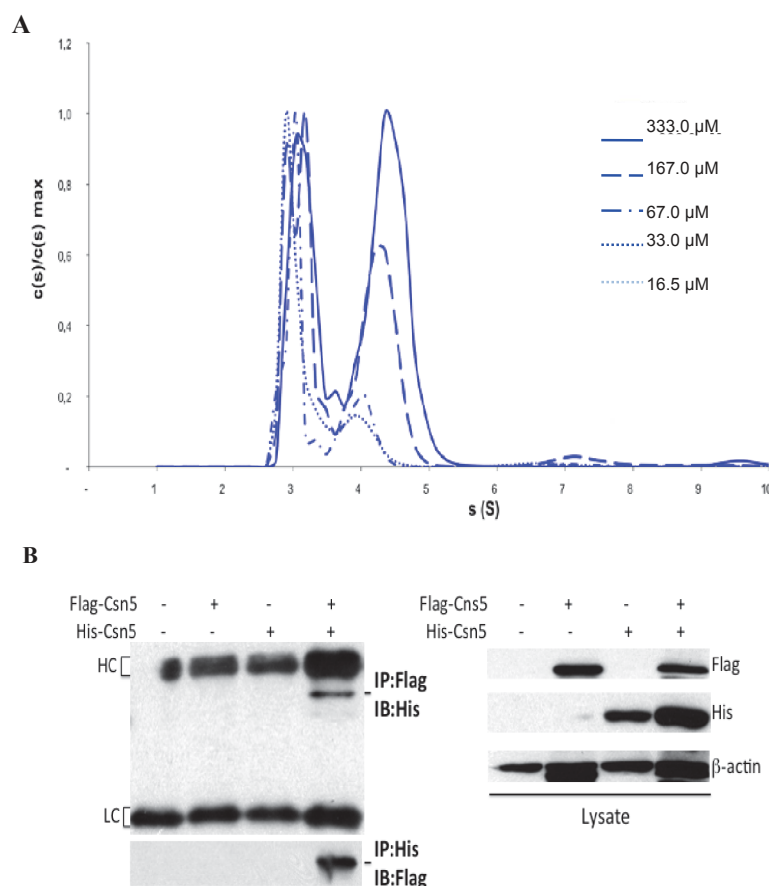
Further to the identification of CSN5 homodimer *in crystallo*, biophysical characterisation of CSN5 dimerisation was initially investigated in solution. The approach used for this was sedimentation velocity analytical ultracentrifugation (SV-AUC) that was carried out by Dr. Christine Ebel (IBS, Grenoble). Processed data were provided to us by our collaborator in Grenoble. This approach was used to separate CSN5 WT into its different oligomeric states in solution. Their sedimentation coefficient  $c(s)$  values, measured in Svedberg units, and the relative amounts of monomer/multimer components in the sample were determined. Data was obtained for different concentrations of CSN5 WT (333.0, 167.0, 67.0, and 33.0  $\mu\text{M}$ ) and the  $c(s)$  values were analysed. The most significant peaks appeared at 3.1 and 4.3 S which most likely correspond to the monomeric and dimeric species of CSN5 WT, respectively (Fig. 3-12. A). Thus the results obtained indicated that, in solution, CSN5 WT is a mixture of monomeric and dimeric species. Larger oligomeric species than the dimeric state of CSN5 were not observed, ruling out the presence of stable tetrameric assemblies. At 333  $\mu\text{M}$ , the equilibrium between monomer and dimer was almost entirely shifted towards dimeric species of CSN5 (Fig. 3-12. A). An evaluated  $K_D$  value for CSN5 dimer (around 700  $\mu\text{M}$ ) reflects a weak association in the tested conditions. Addition of the C-terminal extension (residues 258-334) could affect this value, but the propensity of FL CSN5 to degrade prevented us to evaluate the effect of this region on CSN5 homodimerisation.

As mentioned before, various studies have suggested that, in cells, CSN5 is present in different association states [256, 268]. An *in vitro* study where a mass spectrometry in native conditions approach was used also identified the formation of a total of 35 CSN subcomplexes [128], although it is presently unknown whether this reflects the distribution of CSN subcomplexes in cells. The existence, distribution and function of these different CSN-related species remain unclear. Although CSN5 within the CSN complex is active, its activity state within these potential ‘mini-CSN’ complexes has not probed in depth. Additionally, the sub-cellular distribution of CSN5 might hold a hint as to whether CSN5 has a CSN-independent function. Indeed, initial data had suggested that the CSN complex exists in the nucleus and that the CSN5 subunit was also found in the cytosol. However, recent data show evidence of presence of the CSN in the nucleus, in the cytosol and in chromatin-bound fractions [269]. Most predominantly, the CSN was observed in the cytosol. This study also confirms that, in the cytosol, CSN-free CSN5 form is observed, either in smaller CSN complexes or in the free form. Taken these data together, there is a strong hypothesis suggesting that CSN5 may have a function of its own which does not directly involve the catalytic-related function of the CSN complex. With this in mind, and with our homodimerisation data on CSN5 both in the crystal structure and in solution, the

presence of oligomeric species of CSN5 in human cell extracts was further investigated.

To gain evidence on the presence of CSN5 homodimerisation in human cell extract, co-immunoprecipitation experiments on mammalian cell extracts were performed by our collaborator, Dr. F.-X. Claret. These experiments were performed on the 293T cell line (embryonic kidney cell line), which were transfected with plasmids coding for either Flag- or His-tagged CSN5 or both and incubated for 48 hours. Either anti-Flag or anti-His tag antibodies were used to immunoprecipitate the cell lysate and detection was performed by immunoblotting with anti-His or anti-Flag tag antibodies. Corresponding expression control of Flag- and His-tagged CSN5 was checked by immunoblotting with anti-Flag and anti-His antibodies, respectively. As presented in Figure 3-12B, both His- or Flag tagged FL CSN5 can efficiently immunoprecipitate the alternatively tagged CSN5 subunits, indicating that FL CSN5 forms oligomeric assemblies in cells. These experiments confirmed that the homo-oligomeric arrangement of CSN5 could be observed in a cell extract context, coherently with the crystallographic and biophysical data.

The general tendency of MPN domains to dimerise was highlighted by the fact that many members display dimeric assemblies *in crystallo*, as observed in the MPN domain crystal structures, of *Hs* Rpn8, *Hs* and *Dm* CSN6 and *Af*JAMM. However, in all these cases, only fragments containing the MPN domain were investigated, probably due to difficulty encountered in the purification of FL proteins, as experienced in the case of CSN5. To overcome this, the approach of co-immunoprecipitation of FL CSN5 was used as a method to decipher the argument whether the dimerisation of CSN5 in its FL form can be observed or is a result of the fragment-based approach taken. This allowed, in the context of CSN5, to determine that the dimerisation of CSN5 observed for the fragment 1-257, is unlikely to be a crystallographic artefact formed by the fragment of CSN5 used. Indeed, evidence of oligomerisation was obtained for the FL CSN5 in cell extracts. However, the cellular function of these arrangements of CSN5 was not addressed in our work and remains an interesting aspect to investigate.



**Figure 3-12. Oligomeric arrangement of CSN5.**

**A.** Sedimentation velocity experiments of the WT form of CSN51-257, showing the results of the  $c(s)$  analysis for the protein at 333.0, 167.0, 67.0 and 33.0  $\mu$ M. The dominant peaks at 3.1 and 4.3 Svedberg units correspond to monomeric and dimeric species, respectively. **B.** Co-immunoprecipitation experiments used to biochemically identify CSN5 dimers in cell extracts. Biochemical identification of CSN5 dimer *in vivo* by coimmunoprecipitation experiments. 293T cells were transfected with either Flag- or His-tagged CSN5 or both for 48 h, and then cell lysates were immunoprecipitated (IP) with either anti-Flag (Upper) or anti-His (Lower) tag antibodies and immunoblotted (IB) with anti-His or anti-Flag tag antibodies (Left). The expression of the corresponding pairs of Flag- and His-tagged CSN5 was confirmed by IB using anti-Flag and anti-His antibodies, respectively (Right). IgG heavy chain (HC) and light chain (LC) are indicated.

### 3.5.3 Study of CSN5 homodimerisation: evaluation of the most probable CSN5 dimer in cell extracts

Intriguing results obtained from our investigation of CSN5 homodimerisation guided us to further understand the dynamics of CSN5 homodimer and compare it to its previously studied monomeric state. *In silico*, MD simulations were employed to study the dimerisation of CSN5. The aim of this study was to investigate the dimerisation interfaces of the two CSN5 dimeric forms in further atomic detail, with



the aim to evaluate whether one of these two assemblies is more likely than the other. As described in earlier sections, the crystal structure of CSN5 contains two types of dimers which have a interface surface of comparable properties. As such, the approach described hereafter was used to identify the CSN5 dimer which represent the predominant form in solution and may be of potential biological relevance.

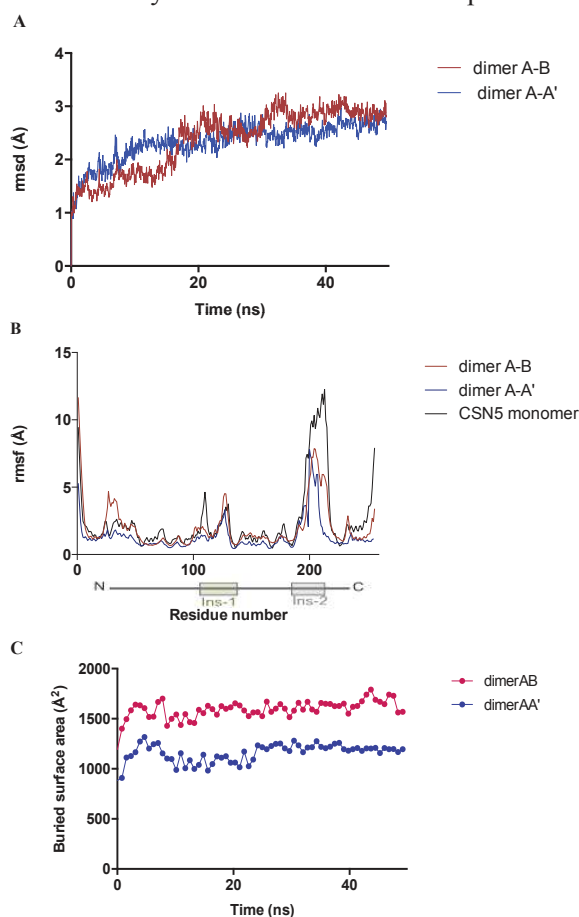
### **3.5.3.1 The dynamics of CSN5 dimers**

For each dimer, A-B and A-A', 50 ns long MD simulations were ran and the outcome trajectories were analysed. A detailed investigation of the behaviour of individual subunits at the atomic level during simulations of the different dimers was carried out. The overall stability of the dimers was measured by estimating the rmsd deviations of each assembly with respect to the initial structures, evaluating changes in hydrogen bonding pattern over the simulations and following the evolution of BSA and ASA values over the course of the simulations. As shown in Fig. 3-13A, the comparative rmsd profiles indicate that each dimer tends to stabilize after 20-30 ns of MD simulation, around a value of  $\sim 2.7$  Å. Analysis of the rmsf of each monomer of CSN5 from the two dimers are compared to the rmsf values obtained previously from our simulation of the monomeric CSN5 (Fig. 3-13B). Overall the expected high fluctuation regions located on the Ins-2 insertion is also observed in both dimers. However the N- and C- terminal regions, which we had previously classified as highly flexible regions, do not display the same behaviours. A difference is seen for dimer A-A', for which the C-terminal part experiences lower rmsf values, reflecting restricted flexibility. This is directly linked to the position of this segment: indeed, the C-terminal region of monomer A is anchored in the active site of monomer A' in A-A' dimer and partially stacked on the monomer B in A-B dimer. Additionally, a common pattern which differs from the monomeric CSN5 is observed in the Ins-1 region of both CSN5 dimers. Whilst, in monomeric CSN5, residues 108 to 113 are highly flexible, in both dimers, the fluctuations in this region are significantly dropped, although the flexibility on the C-terminal helix of the Ins-1 (residues 122 to 130) is maintained in the CSN5 dimers (Fig. 3-13C).

An another way of analysing these dimers is to calculate their interface areas. BSA and ASA analysis over the 50 ns of simulation were plotted to analyse the stability of the dimers (Fig. 3-13. C and D, respectively). Our aim was to use BSA as a criterion to assess protein surface that is not exposed to water and, as such, an increase in the BSA would indicate an increase in the 'compactness' of the binary complex. Results of the analysis indicate that the dimer A-B displays a more extensive buried interface between the 2 subunits, as compared to the A-A' dimer giving an average difference of approximately  $400 \text{ Å}^2$  in BSA. Moreover, whilst the dimer A-A' reaches an equilibrium, with its BSA reaching approximately  $1200 \text{ Å}^2$  in the final 10 ns of the simulation, the BSA of the A-B dimer shows a gradual increase throughout the simulation, starting at values of around  $1400 \text{ Å}^2$  and rising to  $1700 \text{ Å}^2$ , thus indicating

in gain of stability. ASA can be used as a way to validate what we see with BSA calculations, the dimer showing an increase in BSA is expected to have a reduced ASA. Our results clearly show that ASA of dimer A-B is decreasing over the course of the simulation, whilst dimer A-A' remains steady, with a slight tendency to higher ASA values compared to starting values, potentially losing its compactness. Taken together this analysis of the A-B and A-A' CSN5 dimers suggest that the A-B dimer behaves better during the course of the simulations and is more likely to be the stable oligomeric state in solution.

Accurate estimate of thermodynamic properties, for example binding free energy calculations require even more extensive simulations on both dimeric species and thus were not attempted in this work. We used an experimental approach based on the rational mutagenesis of interfacial hydrophobic residues governing the dimer stability to identify the CSN5 dimer which represents the predominant form in solution.



**Figure 3-13. Dynamics of CSN5 homodimers.**

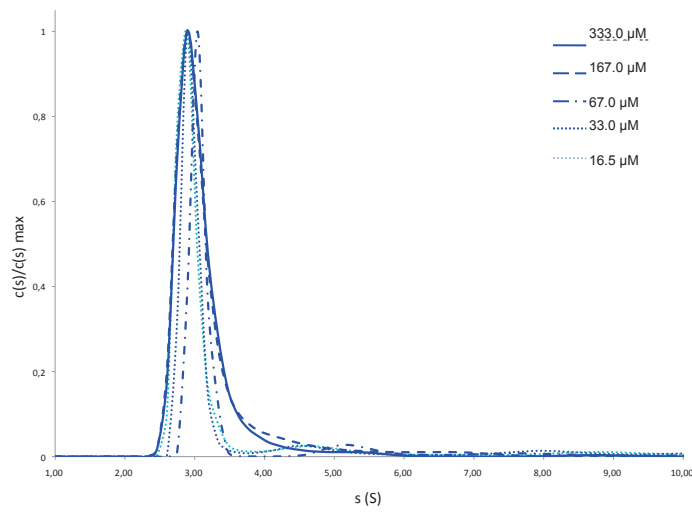
**A.** R.m.s. deviations from the minimized crystal structure of the two CSN5 homodimers during the 50 ns MD simulations. **B.** Fluctuations of each CSN5 monomer from the CSN5 homodimers during MD simulations.  $\alpha$  atom r.m.s. fluctuations of monomeric CSN5 as a function of residue number during the 50 ns MD simulations are shown. **C.** Calculated buried surface area of CSN5 dimers A-B and A-A' over the course of 50 ns simulations. **D.** Calculated accessible surface area of CSN5 dimers A-B and A-A' over the course of 50 ns simulations.

### 3.5.3.2 Validation of CSN5 dimer interface

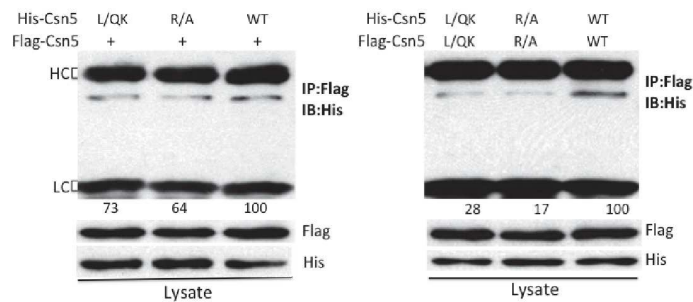
Interfacial mutations were introduced to investigate the disruption of CSN5 dimeric interfaces and to identify a monomeric CSN5 variant. The design of these mutations was guided by data obtained from structural analysis and the MD simulations performed (described in the section above). Targeting the disruption of the A-B dimer, mutations performed include two leucine residues (Leu237 and Leu240) located on the C-terminal helix  $\alpha_6$ , facing the helix  $\alpha_4$  of the second monomer (Fig. 3-11B). The effects of these mutations were evaluated both *in vitro* and in cell extracts using SV-AUC and co-IP, respectively. The CSN5 L237Q/L240K variant was subjected to SV-AUC experiments. Analysis of the sedimentation coefficient at differing protein concentrations of 333.0, 167.0, 67.0, 33.0 and 16.5  $\mu$ M allows to determine a dominate peaks at 2.95 S (Fig. 3-14A; 4-12A for WT CSN5). Analysis of the AUC data on the CSN5 L237Q/L240K variant clearly depicted a drastic decrease in dimeric species of CSN5 in solution. These results confirm the disruption of the A-B dimer in solution and that introduction of the L237Q/L240K substitutions in CSN5 destabilises CSN5 dimerisation, yielding a mostly monomeric protein. Equilibrium displacement towards monomer for the CSN5 L237Q/L240K variant was also clearly visible in the quality of  $^{15}$ N-CSN5 L237Q/L240K HSQC (data not shown; collected, processed and evaluated by Dr. Y. Yang). These results suggest that in solution the variant specifically designed to weaken the A-B dimer assembly displays reduced propensity to dimerise, and thus further supports that the A-B dimer is the predominant assembly in solution, as suggested by our MD simulations analysis.

Results obtained from the investigation of the dimeric interface of CSN5 in solution highlighted that CSN5 in solution consists of a mixture of monomeric and A-B-type dimeric species. This prompted us to confirm the predominance of CSN5 A-B-type dimers in cell extracts. In addition to investigating the CSN5 L237Q/L240K variant, an additional substitution of R129 to an alanine, located on the loop between the helix  $\alpha_4$  and the strand  $\beta_5$  was probed by Dr F.-X. Claret's team. Following the same protocol as previously outlined, mammalian 293T cells were transfected with plasmids with Flag-tag or His-tag sequence, coding for CSN5 L237Q/L240K variant, CSN5 R129 A variant, or CSN5 WT. Cell lysates were subjected to co-IP and IB with anti-Flag and anti-His antibodies, respectively. Comparing the IP of CSN5 WT with the CSN5 interfacial variants indicates that, for both CSN5 L237Q/L240K and CSN5 R129A, there is a decrease in oligomeric assemblies of CSN5 (Fig. 3-14B). Quantification of the bands obtained highlights a significant, 70-80%, reduction in dimer formation. These results confirm the CSN5 A-B dimer is also the preponderant assembly in cell extracts and therefore possibly *in vivo*. On the basis of these results, a study addressing the presence and role of CSN5 homo-dimer in cells would be required.

**A**



**B**



**Figure 3-14. Disrupting the oligomeric state of CSN5.**

**A.** Sedimentation velocity experiments of the L237QL240K (LL/QK) variant form of CSN5 1-257, showing the results of the  $c(s)$  analysis for the protein at 333.0, 167.0, 67.0 and 33.0  $\mu$ M. The dominant peaks at 2.95 and 5.0 Svedberg units correspond to monomeric and, probably, dimeric species, respectively. **B.** CSN5 homodimers are disrupted by the LL/QK and R129A (R/A) mutations. 293T cells were transfected with the CSN5 LL/QK mutation, R/A mutation, or wild type CSN5 for 48 hours and cell lysates were IP with an anti-Flag antibodies and IB with and anti-His antibodies. Cell lysates immunoblotted are shown in the bottom panel. Signal intensities were quantified using ImageJ software.

## Chapter 4: CSN5 activity within the CSN complex

### 4.1 Introduction:

As described in section 1.4 of the Chapter 1, members of the JAMM/MPN<sup>+</sup> family have been found either as single stand-alone multidomain proteins containing deubiquitinating activity or as isopeptidases which are associated in more complex organisations. Interestingly, for the JAMM/MPN<sup>+</sup> protein, Rpn11, CSN5, and BRCC36, the incorporation into their corresponding protein complexes is required for their activation, whilst this does not appear to be the case for AMSH and AMSH-LP [63, 70, 159, 160, 168, 186]. It has been shown in the case of Rpn11 and of BRCC36 that the protein context in which these enzymes are placed influences their activation state. One element that seems to be of particular importance is the presence of a MPN<sup>-</sup> subunit [183]. Indeed in both cases for Rpn11 and BRCC36 and, as we also present here for CSN5, the presence of their respective MPN<sup>-</sup> is sufficient to activate them, as detailed in section 1.4 of Chapter 1. Furthermore, *in vivo* knockdown experiments of MPN<sup>-</sup> subunits give rise to abrogated isopeptidase activity of their corresponding complexes [168]. Whilst the catalytic capacity for MPN<sup>-</sup> subunits in itself has been excluded and their role in MPN<sup>+</sup>/JAMM activation has been established, little was known about how this domain contributes to the enzymatic activity of the complexes. Recent structural and biochemical evidence have suggested that the MPN proteins of these complexes interact directly, forming MPN<sup>+</sup>/MPN<sup>-</sup> heterodimers, within their complexes [67, 158, 160, 168, 259].

In the context of the CSN, the crystal structures of the CSN MPN domain-containing subunits have individually been elucidated. Initially the structure of CSN6 MPN domain, fragment 51-187, from *Dm* at a resolution of 2.5 Å was solved by molecular replacement (PDB code: 4E0Q). This study recapitulated the classical MPN fold commonly described in other proteins containing this domain and further identified the lack of metal binding ability (Fig. 3-1) [64]. However the length of the CSN6 fragment (residues 51 to 184) limited the insight that could be gained regarding its function in the CSN complex. The role of the CSN6 subunit has been heavily accounted for mediating protein-protein interactions. The main reason for this being a general trend of MPN<sup>-</sup> domain-containing proteins which has been depicted in the case of Prp8 and Rpn8 [176]. The importance of protein-protein interactions in CSN6 function was subsequently confirmed, with the N-terminal MPN domain of CSN6 responsible for the interaction with CSN5, whilst C-terminal half involved in its association with the subunits CSN4 and CSN7 [259]. This study suggested that within the CSN complex, the role of CSN6 in keeping the integrity of the holo-complex is essential. Moreover, in *At*, knockout of *csn6* resulted in a global destabilisation of the

CSN complex, accounting for CSN6 essential role in complex assembly [136]. Besides the contribution of CSN6 in the assembly of the CSN complex, it has also been shown that this subunit also contributes to the association of the complex with E3 ligases [47, 70]. However the precise mechanism underlining the function of CSN6 within the CSN complex and its involvement in E3 ligases remain to be fully elucidated.

As described in Chapter 3, molecular determinants of CSN5 activation was established through its active site lock mediated by the salt bridge established between residue R106 and D151 [63]. Here we questioned the activation mechanism of CSN5 within the CSN complex and probed the role of CSN6 in CSN5 activation. Although cryo-EM data of the CSN complex provided evidence that these two subunits are located in close proximity, biochemical confirmation and biophysical characterisation of this binary complex had been lacking [46]. In this Chapter, we investigated the association between the CSN5 and CSN6 subunits and identified that CSN6 is able to activate CSN5, extending the role of CSN6 beyond CSN assembly. Different substrate probes were used to identify the contribution of CSN6 in the activation of CSN5. Moreover the crystal structure elucidation of a human CSN6 fragment (residues 31 to 220) provides further insight to the molecular architecture and aids the understanding of its interaction with CSN5. In addition, the structural information about this subunit helped explain its proposed multi-functional nature. A range of biophysical techniques, used in conjunction with mutational studies was utilised to acquire experimental constraints for the mapping of the interface between the two MPN proteins. Here we provide a reconstructed model of the CSN5-CSN6 heterodimer and characterise its interaction with the Nedd8 part of its substrate. Lastly, given the latest development in the CSN literature, we compare the activation mechanism of the CSN5-CSN6 heterodimer determined with that proposed from the crystal structure of the CSN complex [61].



## **4.2 CSN5 directly binds to CSN6:**

### **4.2.1 Investigation of the direct interaction between CSN5 and CSN6**

The existence in multi-protein complexes of pair of MPN domain-containing proteins, as well as various reports based on yeast two-hybrid, proteomic, non-denaturing mass-spectrometry or electron microscopy studies have brought insights into their dimerisation. To confirm this hypothesis, the association between these two CSN MPN proteins was investigated initially by analytical gel filtration and chemical cross-linking experiments.

An analytical gel filtration assay was designed as a convenient method to check the interaction between the two MPN subunits of the CSN complex. Although several reports had indicated that CSN5 and CSN6 interact, a mass spectrometry-based model of the CSN complex did not support this association [128]. To determine whether these subunits directly associate, CSN6 was titrated into CSN5 to obtain varying ratios of CSN5 R106T:CSN6. As seen in Fig. 4-1A. and B., CSN5 and CSN6 interact and form a gel-filtration stable binary complex. The ratio 1:1 results in a sample that is mostly dimeric.

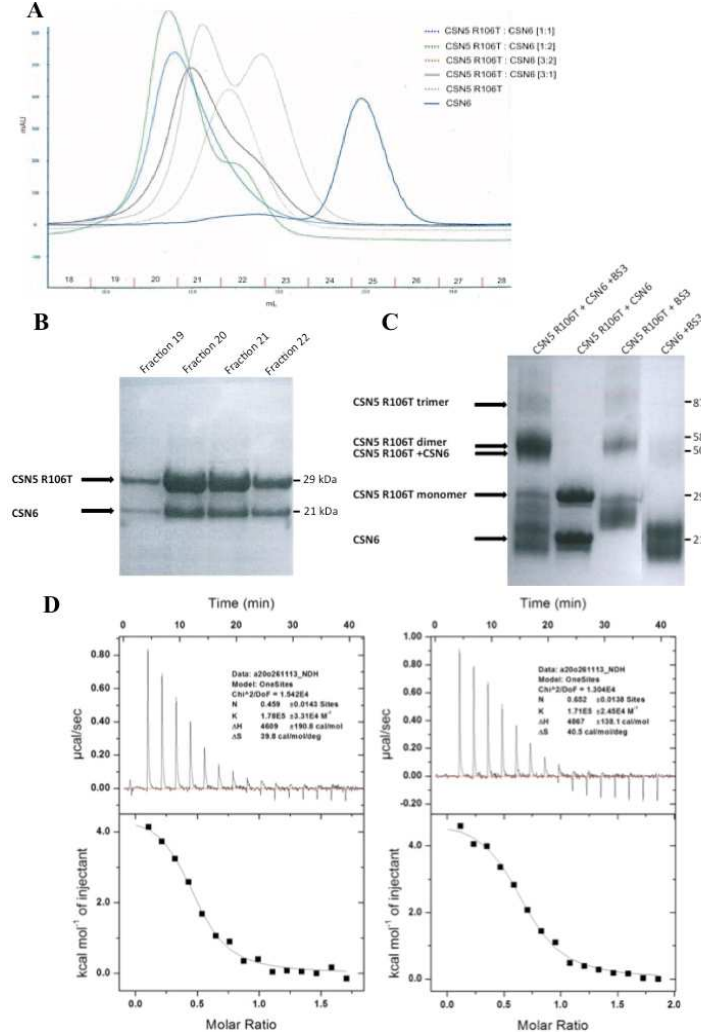
Although the heterodimerisation is clear from the analytical gel filtration experiments, the association between CSN5 and CSN6 was further investigated by chemical cross-linking experiments. The cross-linking agent bis-(sulfosuccinimidyl)suberate (BS3), a bifunctional reagent popular in protein-protein interaction studies, selective for the N-terminal amino group and lysine side chains was used. Denaturing gel electrophoresis of samples after chemical cross-linking showed that the reaction resulted in the production of covalently linked species of different electrophoretic mobility (Fig. 4-1C). In the presence of the cross-linker, the sample containing CSN5 and CSN6, generated three additional protein bands at molecular weights corresponding to the CSN5-CSN6 heterodimer, to the CSN5 homo-dimer and -trimer complexes. The dimeric arrangement of CSN5 in solution was previously discussed in Chapter 3. As controls, CSN5 and CSN6 were individually subjected to the BS3 agent and yielded dimeric, trimeric and higher order CSN5 species, whilst a small band corresponding to the dimeric species of CSN6 was visible.

The results obtained here indicate that the MPN domains of CSN5 and CSN6 form a stable heterodimer complex. The results obtained are also in agreement with reports on the interaction between the MPN proteins Rpn11 and Rpn8 of the proteasome lid, in which the two co-expressed MPN domains form a stable binary complex [158, 160].

### **4.2.2 Affinity of the complex**

The interaction between the MPN domains of CSN5 and CSN6 was further quantified using ITC experiments. Sequential titrations of CSN6 at 1 mM were performed on 100  $\mu$ M of CSN5. Besides confirming the CSN5-CSN6 binary formation, these experiments also revealed a heterodimer  $K_D$  of  $6 \pm 1$   $\mu$ M (Fig. 4-1D). Consistent values regarding the binding properties to CSN6 to the CSN5 constitutively active variant R106T were also obtained, indicating a similar affinity of CSN5 R106T to CSN6 when compared to that of the WT CSN5 (Fig. 4-1D). Our data reveal a comparable affinity of CSN6 for both CSN5 wild type and its conformationally-relaxed variant, indicating that this mutation does not affect CSN5 ability to associate with CSN6.

Here we present the first quantitative analysis of the MPN+/MPN- direct interaction. Previously characterisation of this association has been limited due to challenges associated with the purification of these proteins individually or the co-purification of the two partners [158, 160].



**Figure 4-1. CSN5 and CSN6 form a binary complex.**

**A.** The interaction of CSN5 with CSN6 was monitored by analytical gel filtration. Varying ratios of CSN5 R106T:CSN6 were run on a Superdex 75 10/300 column. **B.** SDS-Page of fractions eluted from the analytical gel filtration column for a ratio 1:1 of CSN5 R106T:CSN6. **C.** Chemical cross-linking experiments using BS3 cross-linker. **D.** The interaction between CSN5 and CSN6 was quantified in ITC titration experiments.

The isothermograms represent the binding of CSN5 (left) and CSN5 R106T variant (right) to CSN6 at 305 K. The  $K_D$  is  $6 \pm 1$   $\mu$ M. Titration raw data: upper panel; integrated data obtained after subtracting the heat of dilution: lower panel. Titration buffer: 20 mM Tris-HCl pH7.5, 50 mM NaCl. Titration: sequential addition of CSN6 (at 1 mM) to CSN5 and CSN5 R106T (at 100  $\mu$ M).

## 4.3 CSN6 stimulates CSN5 isopeptidase activity:

### 4.3.1 CSN5/CSN6 activity on synthetic substrates

The catalytic MPN+ domain-containing proteins that, form part of larger complexes, including the CSN, are known to lack significant enzymatic activity, in their stand-alone form. Various studies on the 19S proteasome lid and the BRCC36 complexes have shown that their MPN- domain-containing proteins are required for expressing the activity of the catalytic subunits. Most recently, the structural studies on *Sc* Rpn11-Rpn8 complex addressed the activity of this heterodimer and showed that this construct, including the MPN domains of these proteins, is in a catalytically active conformation, albeit revealing lower activity compared to the holo-proteasome [158, 160]. It is noteworthy that the specific contribution of Rpn8 in Rpn11 activation was not established in these studies.

Previously we addressed, at the molecular level, the activation mechanism of CSN5 in its stand-alone state (Chapter 3). We showed that CSN5 could be partially activated by the disruption of the active site located salt-bridge between residues R106 and D151. This work highlighted a low affinity of CSN5 for Nedd8 and in particular indicated that its Ins-1 region requires undergoing conformational changes to form a competent recruitment site groove for its substrate [63]. Although insights into the activation of CSN5 had been gained at the atomic level, the activation determinant of CSN5 within the CSN complex remained unknown, i.e. which factors contributes to the structural changes necessary for CSN5 to be in a catalytically primed conformation. The first ‘suspect’ considered was CSN6. We based our work on the three following mechanistic hypothesis: (i) Together with previous literature of MPN-proteins activating MPN+ domain containing enzymes, the direct interaction between CSN5 and CSN6, allowed us to speculate that CSN6 activates CSN5. (ii) The contribution of the substrate could not be neglected, nor could the association of the substrate with other CSN subunits. (iii) It was thus unclear whether the activation of CSN5 is regulated by its incorporation into the CSN or if it is subjected to a substrate induced activation mechanism. Moreover the contributions of the two proposed mechanisms may not be exclusive. We have shown that CSN5 and CSN6 directly interact. Here, we addressed the impact of CSN6 on the isopeptidase activity of CSN5. To do so, we employed the use of different synthetic substrates to conduct enzymatic assays and further characterise the activity using the more complex, physiological neddylated Cul1-Rbx1 substrate.

Initially the activity of CSN5 as a heterodimer with CSN6 was measured against the synthetic substrate, LRGG-AMC. Our results in Chapter 3, revealed that both CSN5 lacks detectable activity on this substrate but its R106T variant displays increased

hydrolytic ability in the tested conditions. In our studies, the addition of CSN6 to the WT and variant forms of CSN5 triggers an unprecedented increase in fluorescence intensity and therefore in catalytic activity for both (Fig. 4-2A). At the highest substrate concentration tested (400  $\mu\text{M}$ ), saturation was not reached. Because of that, full kinetic characterisation could not be obtained and the  $k_{\text{app}}$  values were extrapolated from the slope of the plot  $k_{\text{obs}}$  ( $v_0/E$ ) versus substrate concentration. Comparison of the results obtained reveals a higher catalytic efficiency, of  $0.056 \text{ M}^{-1}\text{s}^{-1}$ , for the CSN5 R106T variant in the presence of CSN6, compared to the CSN5-CSN6, showing 3-fold lower efficiency (Fig. 4-2A). Our results indicate that the loss of the salt bridge between R106 and D151 appears to help the activation mechanism and further a synergetic effect of the presence of CSN6 and of the conformational relaxation of the Ins-1 insertion can be assumed. For comparison, the activity of NEDP1, which contains both deneddylation and pro-Nedd8 processing functions, against LRGG-AMC substrate was investigated. In the cell, the Nedd8 substrate is synthesised as a precursor protein (pro-Nedd8; 81 residues) containing a C-terminal extension,  $-\text{G}^{75}\text{G}^{76}\text{GGLRQ}$ , which is later converted into the mature form (76 residues) via a C-terminal hydrolytic activity resulting in a functional  $-\text{G}^{75}\text{G}^{76}$  C-terminus [38]. The processed Nedd8 is then able to be conjugated to Cullin molecules, which results in active CRLs. In mammals, NEDP1 (also known as DEN1), a 211-residue Nedd8-specific protease, has been shown to be responsible for Nedd8 maturation [53]. This enzyme displays  $k_{\text{app}}$  values of approximately  $4 \text{ M}^{-1}\text{s}^{-1}$ , revealing appreciably more efficient activity compared to the MPN heterodimer investigated (Fig 4-2A).

The activity of CSN5 in the presence of CSN6 was also investigated using the synthetic, AMC-derived substrate, Nedd8-AMC. Previously, in our work, we revealed that, although stand-alone CSN5 lacks catalytic activity, the CSN5 R106T variant is able to hydrolyse the Nedd8-AMC substrate. Comparison of these results with those obtained in the presence of CSN6 highlighted that the activity of both CSN5 and CSN5 R106T are significantly enhanced in the case of the Nedd8 substrate. In these experiments, the saturating concentrations of substrate were not reached. Consequently, full kinetic characterisation could not be carried out and the  $k_{\text{app}}$  was determined to obtain a quantifiable comparison of enzyme efficiency. For the 4 nM CSN5 R106T-CSN6 complex (the concentration of binary complex was calculated using the apparent  $K_D$  values determined by ITC), a strong activity on Nedd8-AMC was detected, giving estimated  $K_M$  and  $k_{\text{cat}}$  values of 18  $\mu\text{M}$  and  $4.3 \text{ s}^{-1}$ , respectively (Fig 4-2B). The same concentration of CSN5 WT-CSN6 complex is unable to be saturated with 6  $\mu\text{M}$  substrate concentration, potentially suggesting a reduced affinity of CSN5 WT/CSN6 for Nedd8. Moreover, comparison of  $k_{\text{app}}$  between these two complexes highlighted that the CSN5 R106T-CSN6 is approximately 4-fold more efficient at hydrolysing Nedd8-AMC, than the CSN5 WT-CSN6 one. The catalytic efficiency of the MPN heterodimer of the CSN can be further compared to data reported on the hydrolysis of Nedd8-AMC by NEDP1. This study showed that this

enzyme contains a significantly stronger catalytic efficiency of about  $41.10^6 \text{ M}^{-1}\text{s}^{-1}$  for the substrate [270]. These results are coherent with ours presented on the hydrolysis of LRGG-AMC substrate by NEDP1, which commonly shows higher efficiency for both Nedd8 and LRGG-AMC substrates.

Furthermore, our results obtained on the enzymatic activity of the CSN5-CSN6 heterodimer are coherent with recent data determined on the kinetic parameters of the 19S proteasome lid Rpn11-Rpn8 dimer. In this work the cleavage of di-Ub was measured determining  $K_M$  and  $k_{cat}$  values ranging from 120 to 299  $\mu\text{M}$  and  $4.8 \times 10^{-3}$  to  $1.2 \times 10^{-2} \text{ s}^{-1}$ , respectively, depending in the type of Ub linkage. The high  $K_M$  values obtained in this work were reasoned as a result of steric hindrance caused by the proximal moiety of the di-Ub. This value is consistent with the  $K_M$  value obtained for the CSN5-CSN6 heterodimer cleavage of NEDD8-AMC in our experiments.

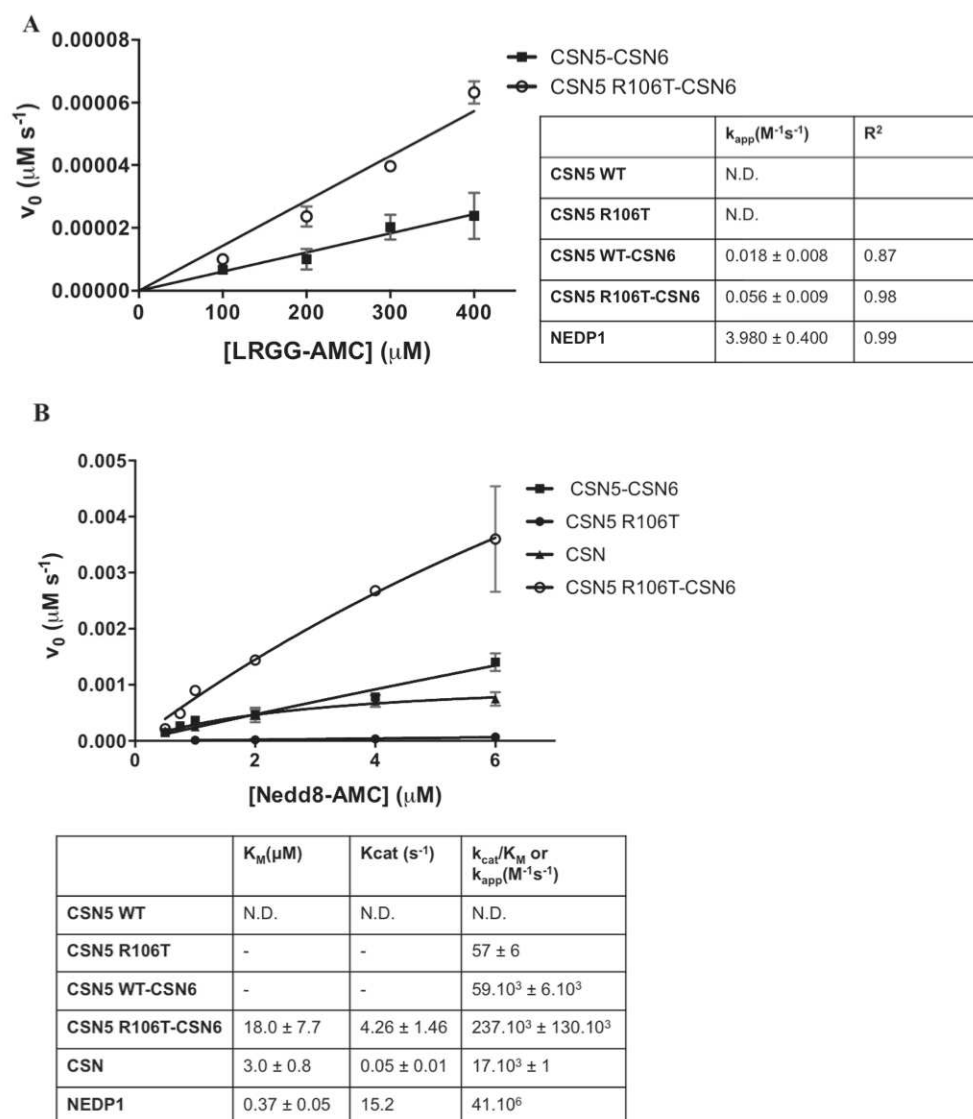
To determine whether the catalytic efficiency determined for these heterodimer complexes is comparable to that observed with the CSN complex, the ability of the holo-enzyme to hydrolyse Nedd8-AMC was investigated. These results revealed  $K_M$  values of 3  $\mu\text{M}$  and  $k_{cat}$  values of  $0.05 \text{ s}^{-1}$  for the CSN (Fig 4-2B). These values suggest at first approximation that the holo-complex could have a higher affinity for the substrate and thus requiring a lower substrate concentration for effective catalysis compared to the CSN5 R106T-CSN6 complex. However the lower  $k_{cat}$  value obtained with the CSN complex dictates that a reduced amount of substrate molecules are turned over per enzyme molecule per second when compared to the CSN5 R106T-CSN6 heterodimer. This indicates a lower overall catalytic efficiency for the CSN complex. These values obtained for the CSN complex suggest that, although its affinity to the Nedd8-AMC substrate is significant, its ability to hydrolyse the substrate is not optimal. The  $k_{cat}$  value of the CSN on Nedd8-AMC is significantly lower than on neddylated Cul1 substrate, as measured by Deshaies and co-workers [72]. This study yielded values of  $K_M$  and  $k_{cat}$  at 212 nM and  $1.1 \text{ s}^{-1}$ , respectively, for CSN on its natural substrate [72], highlighting the contribution of the nature of the substrate to these values. In other terms, these results indicate that neddylated Cul1-Rbx1 is a better substrate for the CSN complex, consistent with various studies [72]. This is likely to result from a stronger affinity between the holo-CSN complex and its natural substrate. Moreover it is possible that neddylated Cullin substrates contribute to the correct positioning of Nedd8-cullin isopeptide bond within CSN5 active site, through extensive interactions between the neddylated CRL substrate and different CSN subunits. In the following section, we further addressed the activity of the CSN5-CSN6 heterodimer on one of its natural substrates and compare it with that of the CSN.

The results presented here in this section clearly demonstrate that CSN6 is able to activate CSN5 from its inactive, stand-alone state, making it a competent enzyme for the hydrolysis of both the LRGG-AMC and Nedd8-AMC substrates. Previous to our

work, whilst the activation of MPN+ domain-containing proteins via their association with their corresponding MPN- domain-containing proteins was reported for both Rpn11 and BRCC36, contradicting data was reported for the relationship between the MPN proteins of the CSN complex. The referred study [143] investigated the significance of the MPN- domain of human CSN6 and a yeast homolog, namely Csi1. Despite sharing substantial sequence homology with the C-terminal domain of human CSN6, the yeast homolog lacks the characteristic MPN- domain. This study suggested that the C-terminus of human CSN6, and not its MPN- domain, is sufficient for the assembly of the holo-complex, forming a fully active organisation. Moreover the absence of the MPN- domain in yeast further encouraged to conclude that the MPN-domain is not a prerequisite for the deneddylase activity of the CSN complex [143]. In addition to lacking a fully corresponding CSN6, subunits CSN3 and CSN8 are also absent in yeast (Table 1-1.) [143]. Furthermore little structural and functional information is available on the activation state of CSN5 in yeast.

Additionally, within our work, using a variety of substrates, we were able to discriminate Nedd8-AMC as a better substrate for the CSN5/CSN6 heterodimer over the minimal LRGG-AMC substrate. Higher affinity for Nedd8 and further interactions with the full Nedd8 substrate may be required for the correct orientation of the isopeptide bond for optimum activity. The similar range of kinetic values obtained for both the CSN and the MPN heterodimer complex, allows us to propose the CSN5-CSN6 complex as the core CSN deneddylase. A major contributor in the activation of the CSN5 subunit appears to be its interaction with CSN6. That being said, the contribution of other subunits has not been addressed here. Previous studies have shown that the CSN5-CSN6 heterodimer forms a distinct appendix, flexibly anchored to the CSN PCI core through their C-terminal regions [46, 61]. As such it is likely that this heterodimer is mostly directly responsible for the catalytic reaction, although other CSN subunits are known to contribute to substrate recruitment in the context of neddylated CRLs.



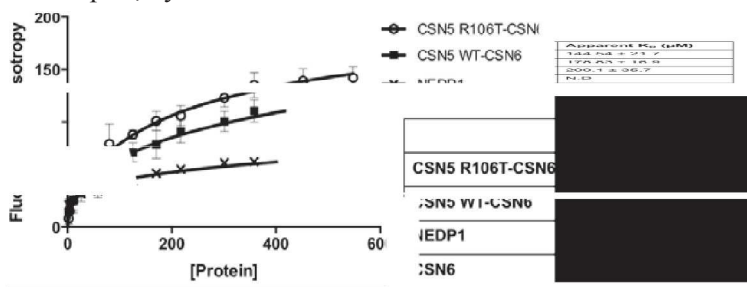


**Figure 4-2. CSN6 activates CSN5 to form the catalytic core of the CSN complex.**  
**A.** The isopeptidase activity of the CSN5-CSN6 complex was tested against the LRGG-AMC substrate. **B.** Kinetic parameters of the isopeptidase activity of the CSN5-CSN6 complex on Nedd8-AMC substrate were determined. Kinetic data for NEDP1 was obtained from [270].

### 4.3.2 The CSN5/CSN6 complex recruits Nedd8

Having both shown that CSN5 in the stand-alone state was not competent for Nedd8 binding and that CSN6 binding to CSN5 triggers its activation, this prompted us to address whether the role of CSN6 in the activation of CSN5 is correlated with an increase in Nedd8 affinity. To explore this, we studied the effect of CSN6 on the interaction of CSN5 with Alexa Fluor 488-labeled Nedd8 by fluorescence anisotropy (Fig. 4-3A). Our results correspond to the binding affinity of the CSN5 WT-CSN6 complex or of the complex containing CSN5 R106T variant, to the Nedd8 substrate. By fitting the data to the Hill equation, apparent  $K_D$  values were determined for each case. Previously we had shown that the conformational relaxation introduced to the Ins-1 region of CSN5 by the R106T variant increases the affinity for Nedd8, as presented in Chapter 3.

Addition of CSN6 to CSN5 WT or to the CSN5 conformational variant R106T revealed in both cases that the presence of CSN6 increases the affinity of the complexes for Alexa Fluor 488-Nedd8 (Fig. 4-3). To clarify whether the increase in affinity observed in the presence of CSN6 is a consequence of CSN6 interacting with Nedd8, its interaction with Alexa Fluor 488-Nedd8 was measured (Fig. 4-3). Here we showed that CSN6 is not interacting with the substrate, although some evidences of very weak interaction had been recorded using NMR (data not shown). The increase in affinity was more pronounced in the case of the CSN5 WT-CSN6 complex where a nearly 2-fold decrease in  $K_D$  was observed, whilst for the CSN5 R106T-CSN6 complex a less significant drop in  $K_D$  was obtained (Fig. 4-3). The same range of affinity observed for the CSN5 R106T stand-alone and, when in complex with CSN6, further suggests that the opening of the Ins-1 region is part of the mechanism involved in forming a competent active site groove that can accommodate Nedd8. Our results indicate that these conformational changes, among which the reorganisation of the Ins-1 region, in the case of CSN5 WT and possibly of CSN5 R106T are promoted, at least in part, by the association with CSN6.



**Figure 4-3. Determination of the dissociation constant of the CSN5-CSN6 interaction with Nedd8.**

The normalised anisotropy of Alexa 488-labelled Nedd8 (4 nM) is plotted as a function of CSN5-CSN6 complex concentration. The  $K_D$  was determined for the CSN5-CSN6 and CSN5 R106T-CSN6 complexes; the binding constant for CSN6 was too low to be determined.

For context, the affinity of NEDP1 for Alexa Fluor 488-Nedd8 was also measured giving values of a similar range to that obtained by the CSN5-CSN6 heterodimer (Fig. 4.3). As a crude generalisation, for proteins containing Ub-binding domains, their association with single Ub or Ub-like proteins has been classified as moderate to weak interactions, typically containing  $K_D$  values ranging from 1-500  $\mu$ M. In biological systems, low range protein-protein interactions are commonly physiologically relevant [264, 271]. Thus the weak binding affinities of Ub-binding domain-containing proteins have been suggested to prevent promiscuous binding of free Ub, which in mammalian cells, exists at a high concentration of approximately 6  $\mu$ M [272]. Examples of affinity studies on MPN domain-containing proteins include that of the pre-mRNA splicing factor, Prp8p, which binds Ub with an affinity of  $\sim$ 380  $\mu$ M, although the relevance of this interaction is presently unknown [265]. More specific for JAMM/MPN+ domain-containing protein, low affinity for the substrate may be considered part of the mechanism of preventing premature activation of their isopeptidase activity. For the heterodimer Rpn11-Rpn8, a tryptophan fluorescence-based assay was reported for its interaction with K48-linked di-Ub revealing a  $K_D$  of  $\sim$ 67  $\mu$ M [160]. We have shown that the affinity of the CSN5-CSN6 dimer is in the reach with  $K_D$  values of other MPN domain containing proteins reported in literature, to date. Inevitably the affinity of CSN5 within the CSN complex for Nedd8 may be stronger, and more so for neddylated Cul1-Rbx1. Factors that may influence the affinity may include the C-terminal helical bundle of the MPN proteins (which are absent in our experimental setup) or further subunit contacts either with CSN5 itself or with the substrate. However correlating with our activity results, the CSN5-CSN6 dimer is in a competent state that can bind and hydrolysis Nedd8-based substrate.

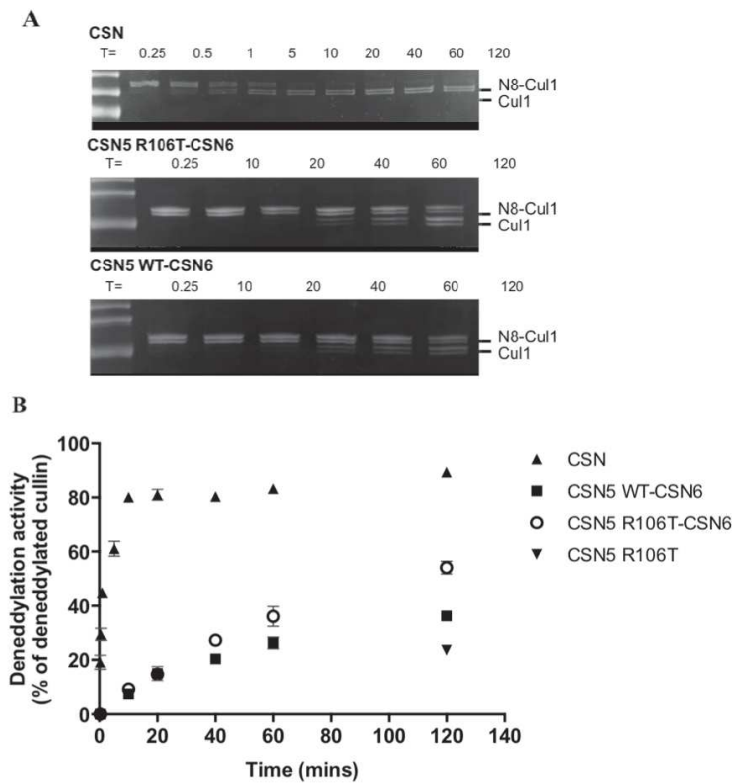
#### 4.3.3 Cullin deneddylation by CSN5-CSN6 complex

Determining that the CSN5-CSN6 heterodimer is a catalytically active complex towards the synthetic substrates, LRGG-AMC and Nedd8-AMC encouraged us to measure its deneddylase activity towards its physiological neddylated Cul1-Rbx1 substrate. To date, the major function of the 8-subunit CSN constitutes the cleavage of the isopeptide bond between the Nedd8 and Cullin proteins [47, 163].

We thus set ourselves to characterise the CSN5-CSN6 deneddylating activity by developing an *in vitro* assay in which the conversion of neddylated Cul1-Rbx1 to Cul1-Rbx1 could be detected and quantitatively measured (Fig. 4-4A). Our results here indicated that, as expected, the CSN at a concentration of 4 nM contains robust activity towards 1  $\mu$ M neddylated Cul1-Rbx1 complex. As previously discussed, the high affinity of CSN towards its physiological substrate was reported in earlier studies where in comparison with our work coherent rates were determined. In both studies the CSN complex hydrolysis more than 80% of its substrate after 10 minutes incubation. In Chapter 3 (Fig. 3-10), we confirmed the lack of deneddylase activity of CSN5 on neddylated Cul1-Rbx1, whilst after 3 hours incubation, the conformational

variant CSN5 R106T showed a modest but significant deneddylase activity. Here in the presence of CSN6, we determined that for both CSN5 WT and its R106T variant (at a 4 nM concentration of the binary complex) against 1  $\mu$ M substrate, approximately 40% and 60%, respectively, of the activity observed for the holo-CSN complex is retrieved (Fig 4-4). Additionally the activity of CSN5 R106T alone on neddylated Cul1-Rbx1 was measured using 900 nM of enzyme and 1  $\mu$ M of substrate, showing a reduced capacity for hydrolysis after 120 min incubation, compared to both CSN5 WT-CSN6 and CSN5 R106T-CSN6 complexes.

Comparison of our results obtained on the hydrolytic activity of the CSN on the synthetic Nedd8-AMC and the physiological neddylated Cul1-Rbx1 substrates further reinforces the already documented contribution of the holo-CSN complex in the efficient binding of neddylated CRLs [46]. A recent electron microscopy study on the CSN in complex with the SCF substrate has shown that the CSN localises to both the functional N- and C- terminal sites of the SCF. In the model proposed, interactions primarily involve the CSN2/CSN5 region forming contacts with Cul1 C-terminus and the CSN1/CSN3 region with the N-terminal substrate receptor module. This indicates a large interaction surface governed on the CSN complex, which can accommodate the SCF. In fact the contribution of CSN5 in this interaction is suggested to be minimal, as this subunit has been shown to be dispensable for the assembly of the remaining seven subunits. This is coherent with the peripheral location of CSN5 within the complex. Moreover, even in the absence of CSN5 it was shown that the remaining subunits could still form a stable complex with SCF, actually suggesting that a CSN5-Nedd8 contact is not required for the CSN-SCF complex formation. This study highly suggested that, although CSN5 harbours the catalytic centre of the CSN, it is itself not fully competent at recognising neddylated SCF substrates in their whole, but requires CSN2 and potentially other CSN subunits to correctly position the Nedd8 moiety in its active site [46].



**Figure 4-4. Deneddylase activity of the CSN5-CSN6 heterodimer.**

**A.** The experiments were performed at 37 °C, in the time intervals indicated on the figure. For comparison CSN (at 5 nM) and CSN5-CSN6 complex (at 4 nM) were incubated with 1 μM substrate. The activity of the heterodimer complex was tested against neddylated Cul1 substrate using a gel shift assay. The bands of neddylated-Cul1 and Cul1 after the experimental run were visualised and quantified. **B.** Plot of deneddylase activity of CSN5-CSN6 binary complex. To obtain a quantifiable value for the deneddylase activity of CSN5 R106T variant, 900 nM of enzyme was incubated with 9 μM substrate for 2 hours at 37 °C.

## 4.4 Human CSN6 MPN adopts a structure that is reminiscent of *Hs* Rpn8:

### 4.4.1 Crystal Structure of CSN6

The recently determined crystal structure of *Dm* CSN6 MPN domain (PDB code: 4E0Q) fragment (comprising residues 51 to 184) corresponds to a portion of the MPN core domain, with various missing fragments difficult to model and lacking its C-terminal region [64]. As a result of this limited structure, our work has aimed to obtain the structure of the human counterpart, preferably of a longer fragment. Three C-terminal truncated constructs of CSN6 were designed on the basis of secondary structure prediction calculated by the server JPred [273]. These three constructs produced protein in expression trials, although the longest fragment was susceptible to protein degradation and the shorter fragment comprising the residues 31 to 220 yielding most robust expression was selected for crystallisation. Additional insights were gained by elucidating the crystal structure of the MPN- core domain (fragment 31 to 220) of human CSN6 by molecular replacement at a resolution of 1.76 Å, where *Hs*Rpn8 was used as the search model (PDB code: 2095; Table 4-1). The crystal belongs to the space group P3<sub>1</sub>21 and contains a single molecule per asymmetric unit. The crystallised human CSN6 MPN structure displays a typical MPN domain arrangement comprising of a central  $\beta$ -sheet made of seven strands decorated with three  $\alpha$ -helices (Fig. 4-5A and B). The human CSN6 core is tightly folded and forms an anti-parallel  $\beta$ -sheet (made of the  $\beta$ 1,  $\beta$ 2,  $\beta$ 3 and  $\beta$ 6 strands) surrounded by the  $\alpha$ 1,  $\alpha$ 2, and  $\alpha$ 3  $\alpha$ -helices. The small anti-parallel  $\beta$ -sheet, consisting of the  $\beta$ 4 and  $\beta$ 5 strands, followed by the  $\alpha$ 2 helix forms the Ins-1 region of human CSN6. The second insertion, Ins-2, in this structure is composed of a  $\beta$ -sheet made of the  $\beta$ 8 and  $\beta$ 9 strands. A C-terminal appendix (comprising the residues 192 to 208) includes a small helical part ( $\alpha$ 4 helix composed on the residues 201 to 207; Fig. 4-5A) that displays signs of flexibility (high B-factor values).



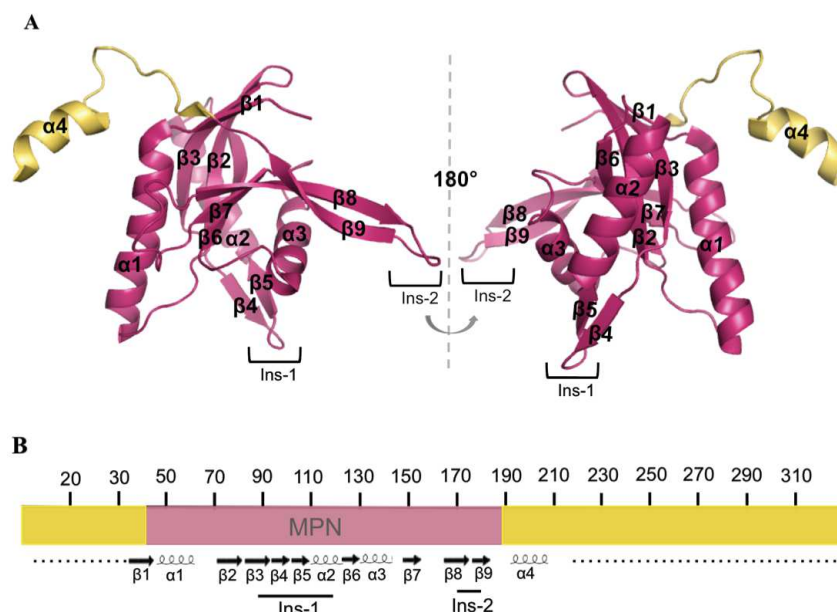
**Table 4-1. Data collection and structure refinement statistics of human CSN6 MPN domain.**

<b>Data collection</b>	<b>CSN6 (residues 31-211)</b>
PDB code	4QFT
Space group	P3 <sub>1</sub> 21
Cell dimensions (Å)	
a	96.84
b	96.84
c	48.33
Maximal resolution (Å)	1.76
Observed reflections	263.696
Unique reflections	26.172
Completeness (%)	100
R <sub>merge</sub> <sup>a</sup>	0.104
Mean I/σ (I)	14.4
Highest resolution bin (Å)	1.86-1.76
Completeness (%)	100
R <sub>merge</sub> <sup>a</sup>	0.492
Mean I/σ (I)	4.9
Refinement	
Protein atoms	1474
Waters	181
Resolution range (Å)	48.00-1.76
R <sub>conv</sub> <sup>b</sup> (%)	15.5
R <sub>free</sub> <sup>c</sup> (%)	20.6
Mean protein B factor (Å <sup>2</sup> )	25.3
Mean solvent B factor (Å <sup>2</sup> )	39.1
Rms deviation bond angles (°)	1.164
Rms deviation bond length (Å)	0.007
Estimated coordinate error (Å)	0.112

<sup>a</sup>R<sub>merge</sub> =  $\sum_h \sum_j |I_{h,j} - \bar{I}_h| / \sum_h \sum_j I_{h,j}$ , where  $I_{h,j}$  is the  $j$ th observation of reflection  $h$ .

<sup>b</sup>R<sub>conv</sub> =  $\sum_h ||F_{oh}| - |F_{ch}|| / \sum_h |F_{oh}|$ , where  $F_{oh}$  and  $F_{ch}$  are the observed and calculated structure factor amplitudes respectively for the reflection  $h$ .

<sup>c</sup>R<sub>free</sub> is equivalent to R<sub>conv</sub> for a 5% subset of reflections not used in the refinement



**Figure 4-5. Crystal structure of human CSN6.**

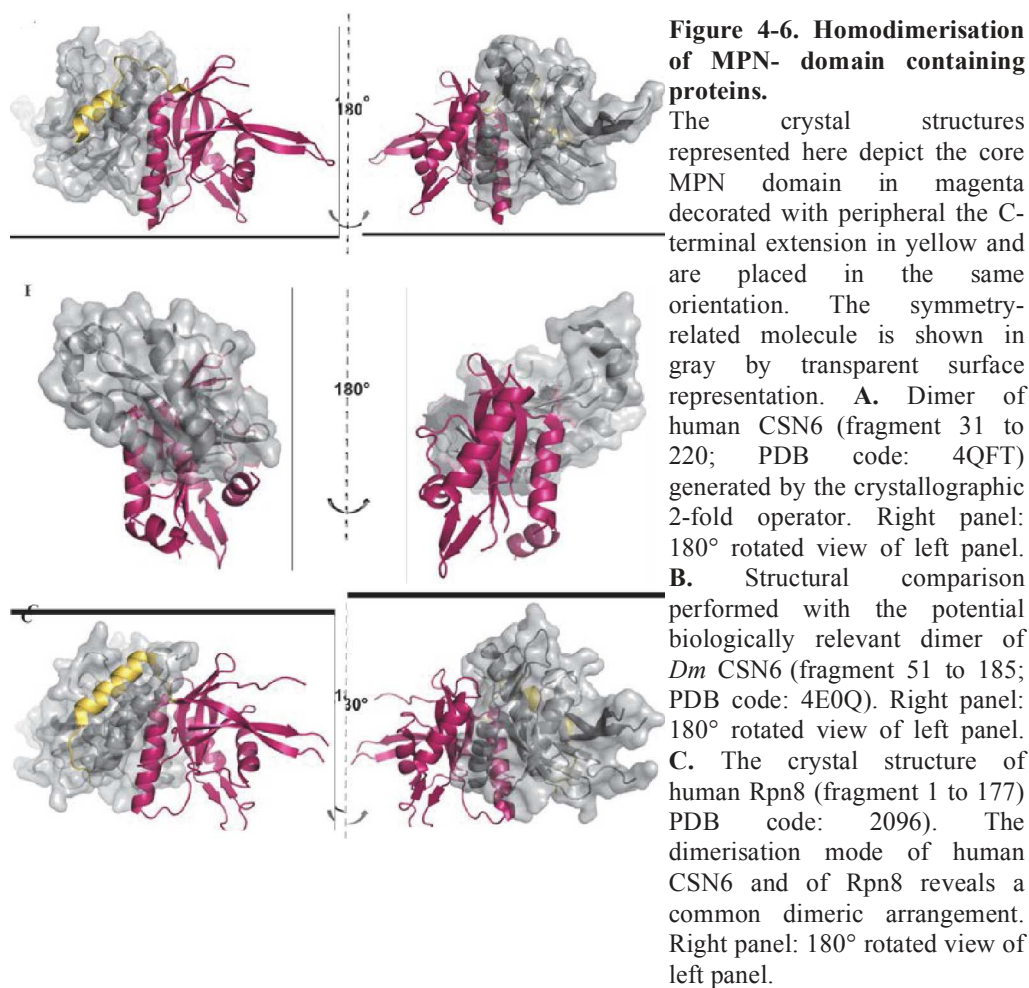
**A.** Crystal structure of human CSN6 (fragment 31 to 220) revealing its core MPN domain (magenta) decorated with a peripheral C-terminal extension (yellow). Secondary structure elements are numbered as indicated. The Ins-1 and Ins-2 insertion have been highlighted. Right panel: 180° rotated view of left panel. **B.** Schematic representation of the CSN6 domain organisation. The MPN domain is delineated in magenta; the N- and C-terminal extensions are yellow. The Ins-1 and Ins-2 insertions are placed with respect to their location in the sequence. The secondary structure elements are indicated. Regions that are either not ordered in the crystal structure or not included in the crystallised fragment are indicated by a black dotted line.

Although a single CSN6 molecule is present in the asymmetric unit, the crystallographic 2-fold symmetry operator generates a dimer stabilised by an  $\alpha$ -helix swapping, via the C-terminal  $\alpha 4$  helix from each of the molecules (Fig. 4-6A). The homodimer interface involves the following secondary structure elements:  $\alpha 1$ ,  $\alpha 2$ ,  $\alpha 4$  on both subunits. According to the PISA Server [210], the interface area for this crystallographic dimer of human CSN6 is 2478 Å<sup>2</sup>. Various other structural examples of MPN domain homodimerisation have previously been demonstrated including the *Dm*CSN6 and the *Hs*Rpn8 dimers [64, 176], illustrated in Fig. 4-6B and C, respectively. However until recently, all the MPN homodimers structurally documented were involving different interfaces.

The *Dm*CSN6 structure identified two dimeric arrangements where the asymmetric unit comprises of the A:B type of dimer with non-crystallographic 2-fold axis, whilst the second B:B type dimer identified is created by crystal packing. The crystal structure of *Hs* Rpn8 also identified two possible dimeric assemblies showing different kinds of dimerisation interfaces: the A:B interface that is contained within the asymmetric unit related by a non-crystallographic 2-fold axis and the A:A dimer representing an interface between symmetrically related subunits. In both structural

studies a biologically relevant dimerisation interface was proposed accounting to the A:B dimer in *Dm*CSN6 and the A:A dimer in Rpn8/Mov34. As shown in Fig. 4-6B and C, the interface of the non-crystallographic dimers of these two subunits differs vastly, however the interface observed in human CSN6 is reminiscent of that described in the A:A dimer of Rpn8/Mov34 crystal structure (Fig. 4-6A and C). The C-terminal helix swapping observed in these two homodimers is a prominent feature. Interestingly, the human CSN6 homodimeric organisation differs to those observed in the *Dm* crystal structure (Fig. 4-6A and B). The difference in the interaction surface observed between human and *Dm* CSN6 structures may be a consequence of the use of a smaller fragment in the case of *Dm* CSN6, resulting in an atypical dimeric arrangement.

The elucidation of the human CSN6 structure allowed us to obtain a more comprehensive view of CSN6 architecture, in particular, regarding the C-terminal part of CSN6 MPN domain and its potential role. Moreover the shared dimeric interface with *Hs* Rpn8 may provide the basis for the JAMM/MPN+ interaction surface with MPN- subunits.



#### 4.4.2 CSN6 conservation and molecular surface properties

A multiple sequence alignment of human CSN6 MPN domain with its homologues depicts a high degree of identity (Fig. 4-7A and B). Mapping the sequence conservation onto our structure of human CSN6 results in a highly dispatched profile. Small conserved patches can be found on the solvent-exposed area, however a large conserved region is identified that is involved in making up the dimer interface. This is suggestive of a potentially important role in CSN6 functions and associations.

Electrostatic surface potential analysis was performed on our structure of human CSN6 revealing that the fragment crystallised preponderantly displays electronegative residues at its surface which is in agreement with its theoretical isoelectric point of 5 (Fig. 4-7C). Regions of strong positive electrostatic potential are located on the C-terminal end of the  $\alpha 2$  helix and the loop between the  $\beta 7$  and  $\beta 8$  strands that together form a solvent exposed groove. The groove corresponds to the site where the canonical JAMM motif is located for JAMM/MPN+ proteins. In addition to the positive charge observed in this region, the groove is also well conserved which could potentially indicate a functional importance associated to protein interactions. Furthermore, the  $\alpha 4$  helix also contains regions of positively charged residues, whilst the Ins-1 and Ins-2 insertions are highly negatively charged.

#### 4.4.3 Comparison of CSN5, an MPN+/JAMM protein with structurally characterised MPN- proteins

As observed with other MPN- domain-containing proteins [64, 176, 274], the human CSN6 structure is void of a canonical JAMM motif and thus lacks a metal or water molecule in the corresponding region. However, comparison of the zinc-binding site of human CSN5 with the equivalent region on CSN6 reveals that several residues forming the JAMM motif in CSN5 are also contained similar in CSN6, suggesting a pseudo metal-binding site. Superimposition of the two subunits identifies that the CSN5 active-site residues, E76, S148, and D151 structurally correspond to residues Q67, D132, and D135, respectively, in CSN6 (Fig 4-7D). The major difference observed is the replacement of the two catalytic histidine residues, H138 and H140, in CSN5 by T126 and G128 in CSN6. This disrupts the zinc coordination site and thus explains the lack of metal-binding ability of CSN6. In CSN6, K153 is seen to occupy the side chain position of H140 present in CSN5 and can potentially restore the negative charge missing from the absence of a histidine, however it is not sufficient for establishing an active catalytic site.

A

```

HsCSN6 -----MAAAAAAATNGTGGSSGMEVDAAV-----VPSVMACG-VTG 37
MmCSN6 -----MAAAAAAGANGSGSSGMEVDAA-----VPSVMASG-VTG 34
XlCSN6 -----MAAAASNG-----NGMEVDVAA-----LPSVMAQG-VTG 28
DmCSN6 MEQMEVDVMSAKPSTSSAAAGSSMAVDKTDQNFPQGNMAAAGTSG 50
SmCSN6 -----MS-----SSS 5
OsCSN6 -----MSAPSDPAVATHFPQAGAAAAAS-----SSS 25
DdCSN6 -----MTDTTTTTTDANKLVLEKSA-----NSS 24
BmCSN6 -----MTEDDSKLAG-----HSEP-----IS 16

HsCSN6 SVSVVALHPLVILNISDHWIRMRSEGRPV-----QVIGALIG 74
MmCSN6 SVSVVALHPLVILNISDHWIRMRSEGRPM-----QVIGALIG 71
XlCSN6 SVTVVALHPLVILNISDHWIRMRSEGRPM-----QVIGALIG 65
DmCSN6 SVTISLHPLVIMNISDHWIRMRSEGRPM-----QVIGALIG 87
SmCSN6 GLTFKLHPLVIVNISDHYTRLMAQAATG-----SAGIPPLRVFGCVIG 50
OsCSN6 GLTFKLHPLVIVNISDHWIRMRSEGRPM-----QVIGALIG 75
DdCSN6 GLEVDLHPLVILNISDHWIRMRSEGRPM-----QVIGALIG 63
BmCSN6 AIAVSLHPLVIMNISDHWIRMRSEGRPM-----QVIGALIG 55

HsCSN6 KQEGRNIEVMSFELLS-HTVEEKIIDKEYYTKEEQKQVFKELFLG 123
MmCSN6 KQEGRNIEVMSFELLS-HTVEEKIIDKEYYTKEEQKQVFKELFLG 120
XlCSN6 KQEGRNIEVMSFELLS-QINDEKITINKEYYTKEEQKQVFKDMFLG 114
DmCSN6 KQEGRNIEVMSFELLS-DVIGDETIVNKDYNNKEEQKQVFKDMLFLG 136
SmCSN6 VQAGRNIEVMSFELLS-DAATD--SLDRPLETKQEQKQVFKDYVFLG 97
OsCSN6 VQAGRNIEVMSFELLS-DPVSG--TLDRPLETKQEQKQVFKDYVFLG 122
DdCSN6 VQAGRNIEVMSFELLS-ATVDRQLVLDIEYLRKK--YEQFLPLDVLG 109
BmCSN6 KQSGRNIEVMSFELLS-KWNNEGKGYAIDADFLSTREQQYQVFKDMLV 105

HsCSN6 WYTTGGPP-DPSDIHVHKQVCEIIESPLFLKLNPM-T-KHTD-LPVSFES 170
MmCSN6 WYTTGGPP-DPSDIHVHKQVCEIIESPLFLKLNPM-T-KHTD-LPVSFES 167
XlCSN6 WYTTGGTP-DPSDIHVHKQVCEIIESPLFLKLNPM-T-KHTD-LPVSFES 161
DmCSN6 WYTTGDNF-TADDIKIQRIIAINECEIMQLNPLS-RSVDHLPLKFLS 184
SmCSN6 WYSIGADI-RESDMVIHKALMDINESPVYLLNPAINPARKELPLSIYES 146
OsCSN6 WYSTGSDV-RDTDMQIHKALMDINESPVYLLNPAINLSPKDLPLSIYES 171
DdCSN6 WYSTGSDV-SKDDILLHKQISEHNSPLYLMLDTS-PKSKDLPLSIYES 157
BmCSN6 WYTTGDTTTPQADWEVHQFSVHESPIHVKLDPKAPPGNK-LPVAIYES 154

HsCSN6 VIDI---INGEATMLFAELTYTLATEEAERIGVDHVARMTAT-GSGENST 216
MmCSN6 VIDI---INGEATMLFAELTYTLATEEAERIGVDHVARMTAT-GSGENST 213
XlCSN6 VIDI---VNGEATMLLAELSYTLATEEAERIGVDHVARMTAT-GSGENST 207
DmCSN6 LIDL---VDGEATMLFVPLTYTLATEEAERIGVDHVARMTAT-ESGKSV 230
SmCSN6 ELHV---IDGVPQLIFVKASYTITVEAERISVDHVAHIKPSDGGSAATQ 193
OsCSN6 ELHV---IDGVPQLIFVRANYTITVEAERISVDHVAHIKPSDGGSAATQ 218
DdCSN6 ELHI---VNDPTTIFVKTPIYKIQTGEAERIGVNHIAKVTPS--GSEGS 202
BmCSN6 VCTTGLSASSPSSVWHSVNVWSLASEEAERIGIEHVAQISTV-STNTVSS 203

HsCSN6 VAEHLIAQHSIAKMLHSRVKLILEYVKASEAGEVPPFNHEILREAYALCHC 266
MmCSN6 VAEHLIAQHSIAKMLHSRVKLILEYVKASEAGEVPPFNHEILREAYALCHC 263
XlCSN6 VAEHLIAQHSIAKMLHSRVKLILEYVRAAGEVPPFNHEILREASALCHC 257
DmCSN6 VAEHLVAQDSIAKMLNTRIKIVLYQIRVDEAGKLRANQELREAYALCHR 280
SmCSN6 LAAHLTGMHSAIAKMLHSRIKIVLHLLAAIQKGADPEHALLRQVSSLVRR 243
OsCSN6 LAAHLTGIHSAIAKMLNTRIRVRIHQVLSMQKGMPLDNLRLQVSSLVRR 268
DdCSN6 LASHLFTMHNAISMLNIRVKALSDYLQVAKKRLPYEQNILRKASSLCNQ 252
BmCSN6 TNKQIAGQLGAIAKMLSRLELIYDYLAVRKGELPRDEAIIREIAQLVRR 253

HsCSN6 LPVLSTDKFKTDFYDQCNDVGLMAYLGTITKTCNTMNQFVNKFNVLYDRQ 316
MmCSN6 LPVLSTDKFKTDFYDQCNDVGLMAYLGTITKTCNTMNQFVNKFNVLYDRQ 313
XlCSN6 LPVLSTDKFKTDFYDQCNDVGLMAYLGTITKTCNTMNQFVNKFNVLYDRQ 307
DmCSN6 LPVMQVPAFQEEFYTCNDVGLISYLGTLTKGCNDMHFVNKFNVLYDRQ 330
SmCSN6 LPAIDSPKFQDDFLMEYNDTLLMTYLATITKCSNTMNEFVEKFTTAYDK- 292
OsCSN6 LPAMESEKFQDDFLMEYNDTLLMTYLAMFTNCSNTMNEFVEKFTTAYDK- 318
DdCSN6 LPTIDTHDFKKSFLQYNDVLLMTYLATITKTSASLNDTIDKYLVSNEKQ 302
BmCSN6 VPMSSDKFLQYKQNCLEVKMTSLVAVLTCTGTLNDLTKMNMVMSVER 303

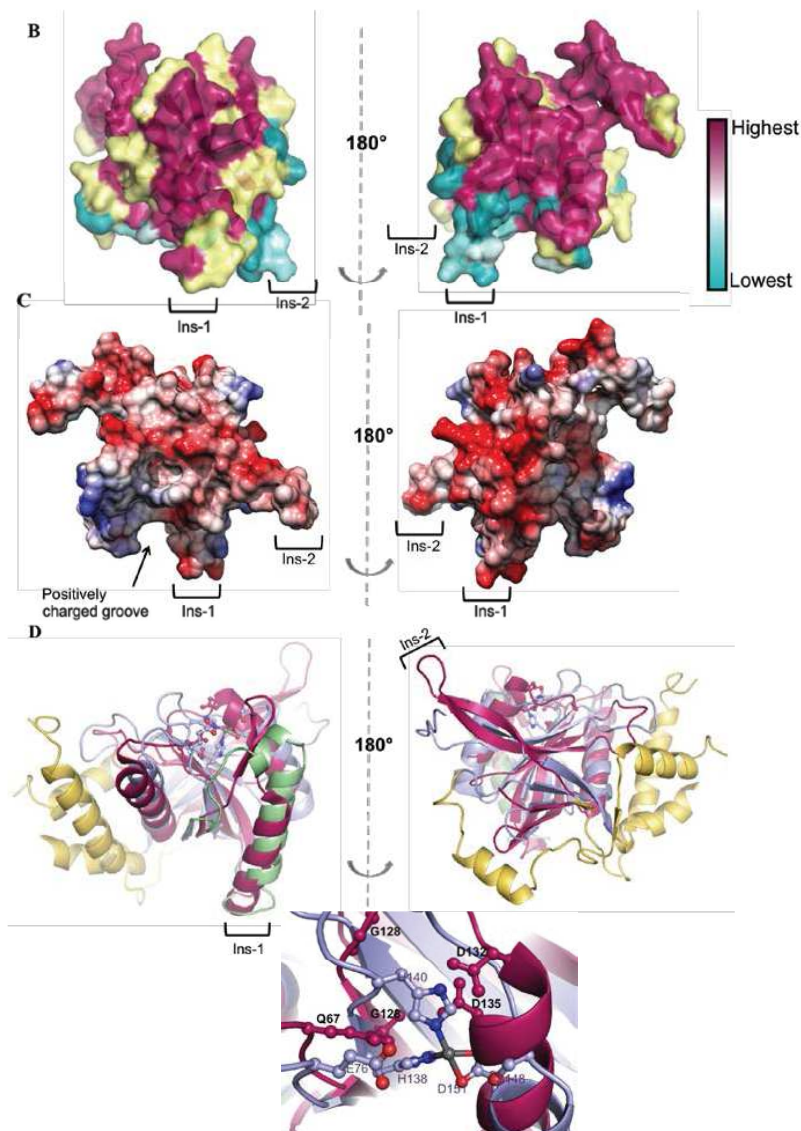
HsCSN6 GIGRRMRGLFF----- 327
MmCSN6 GIGRRMRGLFF----- 324
XlCSN6 GIGRRMRGLFF----- 318
DmCSN6 GSARRMRGLFY----- 341
SmCSN6 QSRRGGRTVFV----- 303
OsCSN6 TARRRGGRGAFM----- 329
DdCSN6 SKRRFYQ----- 309
BmCSN6 QTWLLASGRRARPPFAM 320

```

## Figure 4-7. Biochemical and structural properties of human CSN6.

A. Sequence alignment of CSN6 from eight diverse organisms. CSN6 sequences were retrieved using Blastp with human CSN6 as the search sequence. Subsequent alignment of the nine sequences was performed by Clustal Omega server (default parameters). *Dm*: Uniprot entry: Q9VCY3; *Hs*: Uniprot entry: Q7L5N1; *Mm*: Uniprot entry: Q3UIT2; *Xl* (*Xenopus laevis*): Uniprot entry: Q6NUC2; *Dm*: Uniprot entry: Q9VCY3; *Sm*: (*Schistosoma mansoni*): Uniprot entry: G4V7G0; *Os* (*Oryza sativa*): Uniprot entry: Q6ZKM2; *Bm* (*Bombyx mori*): Uniprot entry: Q2F614. Strictly conserved residues are shown in red. Residues that form the pseudo metal-binding site of CSN6 are indicated with a “\*” sign, in black for residues similar to and in green for those diverging from the canonical JAMM motif. Residues H44, N85, E88 and V115 discussed in the text are in green. The Ins-1 and Ins-2 regions were highlighted in blue boxes, whilst CSN6 residues involved in forming the interface with CSN5 are shown in yellow boxes.





**B.** Mapping sequence conservation on the molecular surface of *HsCSN6*. The sequence conservation was calculated using the ConSurf software and visualized by surface-transparent view. **C.** Electrostatic surface potential of human CSN6 (fragment 31 to 220). Blue and red are positively and negatively charged, respectively. White represents neutral potential. **D.** Superimposition of the human CSN6 (fragment 31 to 220; PDB code: 4QFT) and human CSN5 (fragment 1 to 257 PDB code: 47F0) structures. The Ins-1 and Ins-2 regions are labelled. A close-up representing CSN5 JAMM motif of CSN5 superimposed with the pseudo metal-binding of CSN6. Residues in violet are belonging to CSN5, whilst those in magenta are CSN6 residues.

## 4.5 Structural characterisation of the CSN5/CSN6 complex:

The heterodimerisation of MPN domain-containing proteins have long been anticipated with evidence coming from their X-ray structures and from EM data both for the CSN complex and the 19S proteasome lid. Among the crystal structures of MPN domains, characterised to date, including those of *AfJAMM* [65], *AMSH-LP* [186]; *AMSH* [275]; *HsRpn8* [158, 160, 176], *DmCSN6* [64], *HsCSN6* [276], *Prp8* [274] and *CSN5* [63], available to date, only few (*AMSH*; *AMSH-LP*; *Prp8*) do not display evidence of dimerisation (Table 4-2). For most, excluding *AfJAMM*, dimeric species of MPN domain-containing proteins was also identified in solution. Interestingly, many of these MPN proteins display in their crystal structure two different kinds of dimers, one within the asymmetric unit and the other created by symmetry operators. Relying on differences in solvation free energies for interface formation between the two dimer interfaces, calculated by PISA [210], detected in *HsRpn8* and *DmCSN6*, allowed to distinguish the biologically relevant dimerisation interface. For *HsRpn8*, the biological dimer was suggested to be the dimer resulting from crystal packing, as it is also the case for human *CSN6*, whilst the asymmetric unit potentially forms the physiologically relevant dimer of *DmCSN6*.

**Table 4-2 Summary of the MPN-MPN contacts (in the context of homo- or hetero-dimers) in MPN crystal structures.**

Interface surface areas were computed by the PISA server [210]. (?): no additional experiments have been reported to establish the oligomeric state. (\*): although it has been shown in cellular context that the full-length protein does not homo-dimerise.

Protein	PDB code	Surface area (Å <sup>2</sup> )	Potential homo-dimer
<i>AfJAMM</i>	1OI0	720	X (?)
	1R5X	730	
<i>AMSH</i>	3RZU	700	
	3RZV	615	
<i>AMSH-LP</i>	2ZNR	550	
	2ZNV	260	
<i>HsCSN5</i>	4F7O	1055 974	X
<i>HsCSN6</i>	4D18	3798 (with CSN5)	heterodimer ?
	4QFT	2478	
<i>DmCSN6</i>	4EQ0	965 694	X (*)
<i>HsRpn8</i>	2O95	2300/600	X (?)
	2O96	1100/750	X (?)
<i>ScRpn8/ScRpn11</i>	4O8X	1981	heterodimer
	4OCL	1938	heterodimer
<i>Prp8</i>	2P87	600	

As detailed in Chapter 3, the two types of homodimers observed in the crystal structure of *CSN5* showed comparable values for interface analysis making their



differentiation more challenging. In our work, we determined that the A-B dimer within the asymmetric unit, involving the helices  $\alpha 4$  and  $\alpha 6$  secondary structure elements, represents the most probable biological interface of CSN5. The particular interest in the studies of MPN homodimer interfaces is that it may provide the basis for the MPN<sup>+</sup>-MPN<sup>-</sup> heterodimer complexes. Although dimerisation seems to be a recurrent principle of MPN domain containing proteins, a considerable variation in the mode of dimerisation is exhibited. As the time at which we started investigating the interaction between CSN5 and CSN6, it had remained unclear which of the observed interactions, if any, were biologically relevant.

To date EM studies of the 19S proteasome lid [175, 258], the CSN complex [46, 266, 277] and of the eIF3 complex [267] have been reported. The general similarity of the organisation of these three complexes suggested that their MPN domain-containing subunits would display alike assemblies. In these three complexes, the MPN domain-containing proteins are found in close proximity. The dimeric nature of the MPN domains of these proteins and their location within their complexes, together, suggest that these pairs of subunits could interact through their MPN domains. This is supported by our findings that CSN5 and CSN6 associate directly. Furthermore, recently, whilst our work was in progress, the first heterodimer arrangement of an MPN<sup>+</sup> with an MPN<sup>-</sup> domain-containing protein was reported in the crystal structure of the yeast Rpn11-Rpn8 binary complex of the 19S proteasome lid [158, 160]. These two studies, confirming our at-the-time-unpublished findings reported that this heterodimer adopts a dimeric arrangement oriented in a manner reminiscent of the *Hs*Rpn8 homodimer structure and of our *Hs*CSN6 homodimer structure [158, 160].

The identification that CSN5 and CSN6 directly interact encouraged us to investigate the interface between these two subunits that mediates their heterodimer arrangement. And we therefore aimed to obtain high-resolution structural and mechanistic insights on this binary complex, which could help to unravel the unknown activation of CSN5 within the CSN complex. Here we present the two approaches taken to accomplish this; interface investigation by NMR followed by sequence conservation analysis and mutagenesis studies, and high-resolution structure of the binary complex by X-ray crystallography.

#### **4.5.1 Crystallisation of the heterodimer complex**

Following our findings that CSN5 and CSN6 form a gel filtration stable binary complex and further determining a  $K_D$  of 6  $\mu$ M prompted us to employ X-ray crystallography to gather structural information on the CSN heterodimer complex. Each subunit (CSN5 fragment 1-257; CSN6 fragment 31-220) was expressed and purified individually and were then incubated together in a mixture on ice for one hour before being concentrated. Initial commercial crystallisation plates (Classic I, PACT) were set up to determine an optimum concentration of 7 mg/mL for the

complex. We then used this concentration to screen several other conditions. Despite the vast number of conditions screened both at 18°C and 4°C, only one hit was found in the Pro-complex screen (composition: 0.1 M MgCl<sub>2</sub>, 0.1 M HEPES pH7, 15% PEG 4000) at 18°C after one week producing small bi-pyramidal-shaped crystals. Extensive refinement of these small crystals was carried out, to obtain sufficient sized crystals for mounting and collecting diffraction data, by screening around the hit conditions. Manually prepared screens were set up ranging in pH, from 5 to 8, and in concentration of precipitating agent, from 5% to 30%. Regardless our efforts to reproduce and optimise the crystals obtained from the commercial screen, further crystals were not obtained. We thus employed microseeding technique to improve our crystal quality and size. The principle of microseeding is to control crystallisation by avoiding the nucleation zone. The small crystals obtained in the hit solution were fished, crushed and suspended in a slurry of mother liquor. The microcrystals created were then used to set up crystallisation screens. This approach identified one hit condition in the JSCG+ screen. The resulting crystal was mounted for data collection, but did not show signs of diffraction on the ESRF ID29 beamline.

Although difficult to rationalise, another often mentioned prerequisite for successful crystallisation is the absence of unstructured, highly flexible regions within a protein, including extended loops or tails. Specifically for CSN5, although previously the crystal structure was obtained, we identified the highly mobile Ins-2 region for instance that is disordered in the crystal. We thus removed such regions from the CSN5-CSN6 complex by carrying out limited proteolysis (using the limited proteolysis kit implemented by Dr. François Hoh) in order to search for a stable fragment. Three dilutions of each protease were tested and we found that trypsin (1000:1 protein:trypsin ratio) gave most promising results by producing one major protein band corresponding to a smaller sized CSN5-CSN6 complex. We further used these conditions to perform *in situ* limited proteolysis for crystallisation. Trypsin solution was incubated with the protein complex mixture for thirty minutes prior to setting up the crystallisation plates at 18°C. Despite identifying two hit conditions giving small protein crystals, they were not reproducible in subsequent refinement screens. However, the initial crystals obtained from the commercial screen were mounted and exposed to X-ray. The crystal was unable to provide a diffraction pattern of the complex, further confirming the protein nature of these poorly ordered crystals.

In spite of extensive trials, the crystallisation of human CSN5-CSN6 heterodimer was not successful. Within our work, the poor solubility, highlighted by the heavy precipitation of the protein complex solution in many conditions and the high flexibility of the CSN5-CSN6 were evident observations. These two physical principles are also major factors that may influence crystallisation. However, by performing limited proteolysis and still not obtaining crystals may suggest that it is not a matter of flexibility but rather the poor solubility and/or overall dynamic of the complex that accounts to the limited crystals obtained. Whilst crystallisation trials

were in progress, in parallel other biophysical techniques were employed to investigate the biochemical characterisation of the CSN5-CSN6 complex.

#### 4.5.2 Mapping the interface of the CSN5/CSN6 complex

MPN domains tend to form homodimers through various interfaces and this may reflect the interface of the MPN heterodimer arrangement. Moreover, reports of the crystal structure of the Rpn11-Rpn8 binary complex from yeast provided a model system to compare these MPN heterodimers. It is important to stress that virtually all the work described here was carried out before the publication of the Rpn11-Rpn8 structural studies. As CSN5-CSN6 failed to crystallise readily, we employed NMR spectroscopy, under the direction of Dr. A. Padilla together with other biophysical methods, as well as mutagenesis experiments to identify the mode of interaction present in the CSN5-CSN6 heterodimer.

NMR chemical shifts of protein residues (through  $^{15}\text{N}$  and  $^1\text{H}$  signals) are highly sensitive to changes in the local environment. Therefore chemical shift perturbations can be used to map intermolecular interfaces of protein complexes [278]. NMR HSQC titration experiments of CSN5 with  $^{15}\text{N}$ -uniformly labelled CSN6 were performed. Prior to performing titration experiments, a 3D  $^1\text{H}$ - $^{15}\text{N}$ -NOESY-HSQC spectrum of CSN6 (600  $\mu\text{M}$ ) was acquired for the assignment of the side chains of this subunit by Dr. Yinshan Yang. Assignments were obtained for  $^{15}\text{N}$ , HN and Ha atoms producing an overall assignment completeness for  $^1\text{H}$  resonances at 79 % (Fig. 4-8A). Proline residues, as well as the G40-L42/L59-L60 stretches and the residues D71 and H130 were not assigned. Additionally, a temperature screen was performed to determine an optimum temperature to conduct HSQC experiments of CSN6. Indeed this was important to obtain exploitable data as the heterodimer complex reaches a size of 50 kDa. We determined that at 308K the most useful spectrum was obtained and this temperature was used in the following experiments.

Titration of  $^{15}\text{N}$ -labeled CSN6 at 100  $\mu\text{M}$  with increased concentration of unlabeled CSN5 (up to 200  $\mu\text{M}$ ) caused large shift changes of both amide proton and nitrogen resonances of several CSN6 residues (Fig 4-8B). The vast amounts of chemical shifts indicate potential residues involved in the interaction between the two proteins (although it could also be due to conformational changes), whilst the peak broadening and disappearance could suggest conformational changes occurring on CSN6 upon binding to CSN5. Additionally, the size of the CSN5-CSN6 complex of calculated mass at 50 kDa is on the large size front of proteins that can be investigated by NMR. This resulted in a substantial broadening of the peaks and introduced a further difficulty to unambiguously ascribe the CSN6 residues perturbed as a result of the interaction with CSN5. Upon addition of CSN5, the observed gradual loss in intensity of the signals in the CSN6 spectrum is suggestive of a large complex formation and of an intermediate chemical exchange binding, coherent with the lower  $\mu\text{M}$  dissociation

constant determined by ITC ( $6 \pm 1 \mu\text{M}$ ). From the spectrum obtained we can clearly depict an interaction occurring but we could also hypothesize that the drastic changes, including a large number of resonances shifting or reducing in intensity, and several being entirely absent, observed may be a result of conformational changes induced on CSN6, upon CSN5 binding.

Moreover, the HSQC experiments highlighted the fact that the CSN5-CSN6 complex has limited solubility in the tested conditions, as suggested by the precipitation of the sample after overnight data collection, although this is not an atypical sample behaviour. This observation is coherent with the crystallization trials and explains the reason why we were not able to carry out a full HSQC titration of CSN5. Attempts to troubleshoot this physical limitation of solubility was investigated by applying changes in temperature, pH and protein concentration, however these changes affected the spectrum recorded for CSN6 alone and better conditions were not determined. In summary at this stage, we were confronted to a situation in which, although HSQC spectrum changes could be observed, we could not absolutely confirm that all the changes were due to short range residue-residue interaction, and we could not exclude long range effects. Consequently, as a result of the challenges encountered to specifically assigning chemical shifts occurring upon addition of CSN5 to  $^{15}\text{N}$ -CSN6, a reverse approach was taken where the  $^{15}\text{N}$ -CSN6 HSQC was analysed by extracting the residues that remained unperturbed upon addition of CSN5. Data from this approach highlighted residues that are not primarily involved in the interaction (Table 4-3; Fig. 4-8C). Some of these residues mapped on non-exposed regions and are located in the core of CSN6, while others mapped on the outer surface. Highlighting these residues on our human CSN6 structure molecular surface defined two exposed regions, one which groups most of the unperturbed residues and the other which mostly lacks unperturbed residues. Solely from our NMR data however, this latter surface could not be defined as the surface implicated in the interaction with CSN5. Further data were needed to confirm the involvement of this surface in the CSN5/CSN6 association.

To further validate our NMR data, the sequence conservation pattern of CSN6 was analysed. Various studies have determined that functionally important surfaces were enriched in conserved residues, in comparison with the remaining accessible surface [279, 280]. Our rationale was to determine whether the CSN6 surface identified by NMR which, may be the CSN5 binding interface, contains a larger population of conserved residues. However, CSN6 has been attributed to contain various interaction partners and thus conserved regions outside of its interaction interface with CSN5 were expected. Indeed mapping sequence conservation on the human CSN6 structure highlighted a surface of highly conserved surface residues corresponding to the same region accommodating the residues which according to our NMR data may potentially be implicated in the interaction with CSN5 (Fig. 4-8C). In addition the surface containing unperturbed residues contains low sequence conservation whilst

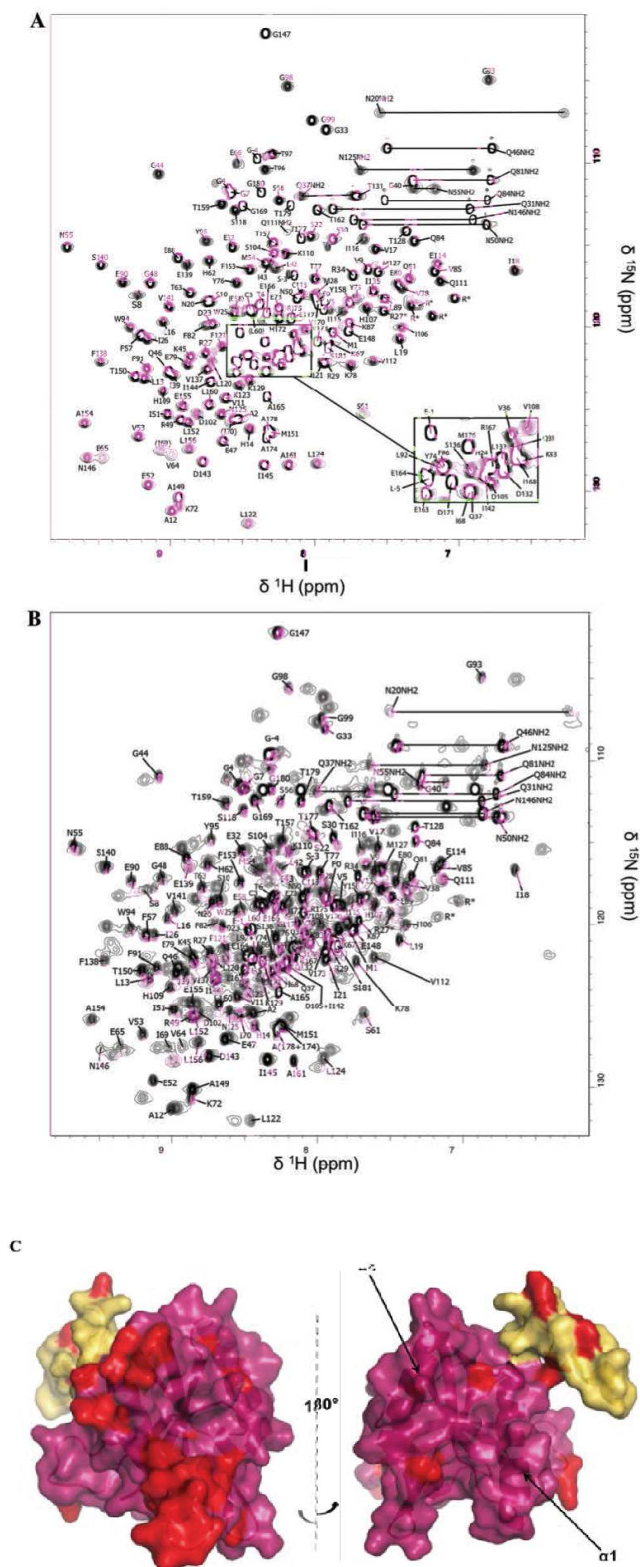
other conserved patches were determined possibly accounting for the interaction of CSN6 with other binding partners (Fig. 4-7B).

**Table 4-3. CSN6 residues found unperturbed, upon binding with CSN5 as evaluated by NMR HSQC.**

Residues marked with an asterisk (\*) are not present in human CSN6 crystal structure.

C33*	M84	V171	L182
G34*	L90	I172	L186
V35*	K97	D173	T187
T36*	E109	I174	T189
G37*	L119	I175	E196
S40	H139	N176	G199
G78	K140	G177	V200
R79	F151	A179	H202
N80	K159	T180	T209*

Moreover, structural analysis depicted that the surface highlighted both by NMR and sequence conservation analysis is also the one governing the homodimerisation of both *HsRpn8* and *HsCSN6* structures. It is noteworthy that this also meant that we could possibly deduce the interaction surface of CSN5 with CSN6. The data presented here suggests that the surface of CSN6, which is engaged in the binding to CSN5, is mostly likely to be comprised of interactions mediated from helices  $\alpha 1$  and  $\alpha 2$  and the surface in between these two (Fig. 4-5A; 4-6A). Our findings were further validated by the crystal structure of the CSN5-CSN6 homolog complex, Rpn11-Rpn8, where the corresponding helices on Rpn8 (also  $\alpha 1$  and  $\alpha 2$ ) are directly implicated in forming two distinct parts of the complex interface. Next, using the binding surface of CSN6 determined, specific variants of CSN6 were designed to probe the interaction with CSN5. The functional relevance of these interactions was validated by activity and binding assays.



**Figure 4-8. Elucidation of the interface between the CSN5-CSN6 heterodimer.**

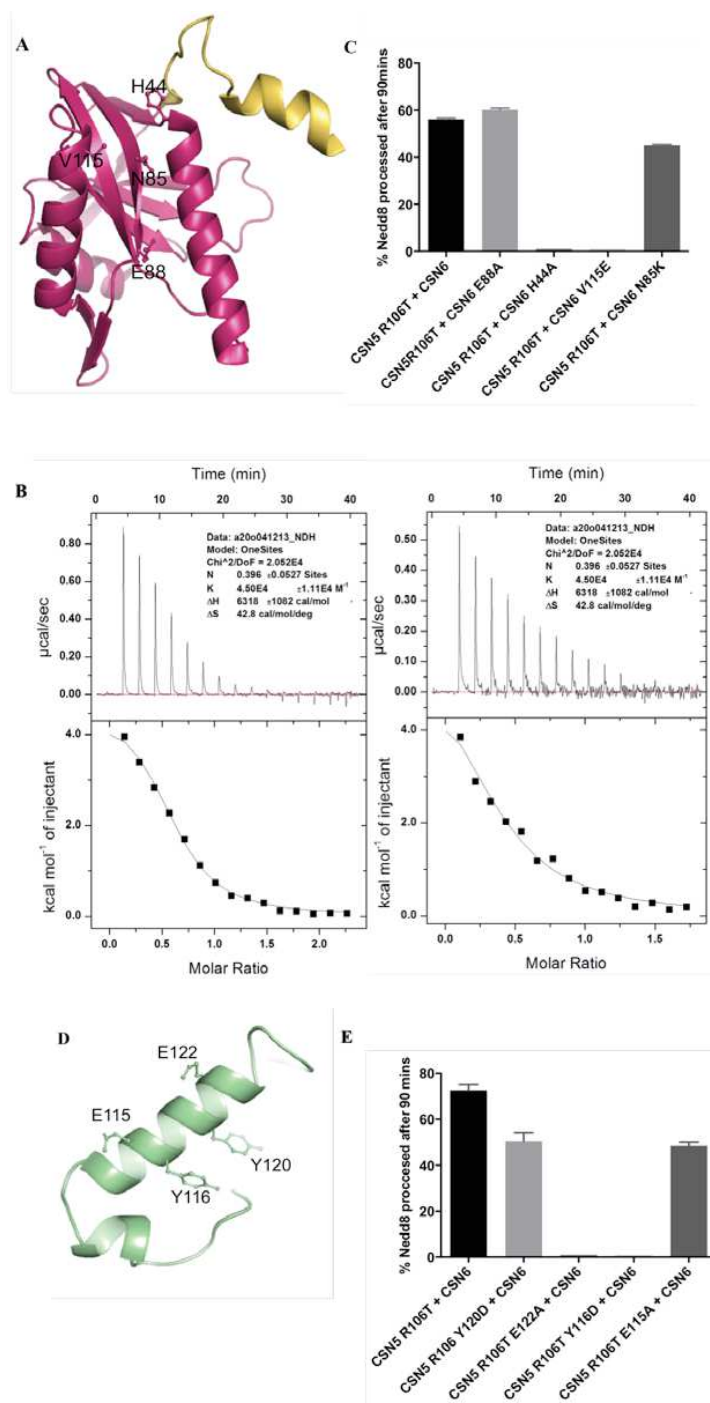
**A.** The assigned  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of human CSN6. **B.** The overlaid 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra highlighting the difference between free human CSN6 (magenta) and CSN5 bound CSN6 (black). **C.** Mapping interfacial data obtained from NMR on the structure of human CSN6. Residues of CSN6 which have not been perturbed after the addition of CSN5 are highlighted in red and illustrated with a surface-transparent view. Helices  $\alpha 1$  and  $\alpha 2$  are highlighted as potentially forming the interface between the CSN5-CSN6 complex. Right panel: 180° rotated view of left Panel.



#### **4.5.3 Interfacial mutations on the CSN5-CSN6 complex**

Further validation of the interaction surface of CSN6 in the complex that it forms with CSN5 was addressed by conceiving and obtaining interfacial variants of CSN6 by site-directed mutagenesis. Here we aimed to probe the potential interaction surface and to confirm the binding interface identified. Chosen due to their high conservation across CSN6 orthologues and their location on the presumptive interaction surface, two CSN6 variants were designed. The nature of the substituting amino-acid was chosen to be different from the original position: for H44 located on the helix  $\alpha 1$ , from a bulky and polar histidine amino-acid to a small and hydrophobic amino-acid, alanine (H44A) and for Val115 located on the helix  $\alpha 2$ , from a small and hydrophobic to a large and charged amino acid, glutamic acid (V115E) (Fig. 4-9A).





**Figure 4-9. Mutations performed to validate the CSN5-CSN6 heterodimer interface.**

**A.** The putative interfacial residues mutated mapped on the structure of human CSN6. **B.** The effect of the interfacial mutants, CSN6 H44A and CSN6 V115E were evaluated by ITC titration experiments. The isothermogram represents the binding of CSN5 R106T variant to CSN6 H44A (left panel) and CSN6 V115E (right panel) at 305K. The raw data of the titration of CSN5 R106T with the CSN6 variants are shown in the upper panel and the integrated data obtained after subtracting the heat of dilution are shown in the lower panel. The titrations were performed in 20 mM Tris-HCl pH 7.5, 50 mM NaCl. **C.** The effect of the CSN6 interfacial mutants on the pro-Nedd8 processing activity of the complex was tested and compared to the WT activity. The pro-Nedd8 variants by their ability to be able to be processed by CSN5 R106T-CSN6 complex. The plots show the percentage of pro-Nedd8 processed after 90 mins incubation at 37°C. **D.** The Ins-1 segment of CSN5 with the residues that have been mutated shown in ball-and-stick mode. **E.** The CSN5 R106T variants are evaluated by pro-Nedd8 processing activity. The plots show the percentage of pro-Nedd8 processed after 90 mins incubation at 37°C.

The effect of these mutations on the association with CSN5 was monitored by ITC experiments on CSN5 and the H44A and V115E variants of CSN6. Both CSN6 variants bind to CSN5 with a moderately lower affinity with a  $K_D$  of  $9.2 \pm 1.0 \mu\text{M}$  for H44A and  $29.0 \pm 2.1 \mu\text{M}$  for V115E (Fig. 4-9B). Compared to the WT CSN6 showing a  $K_D$  of  $6 \pm 1 \mu\text{M}$  (Fig. 4-1D), admittedly but also not surprisingly, the effect impaired by the variants is modest. Following our experimentally-based hypothesis that the CSN5-CSN6 complex is mediated in a similar manner as that of the *Hs* Rpn8 homodimer and the *Sc*Rpn11-Rpn8 heterodimer, it is perhaps not surprising that a large interface is not entirely disrupted by one amino-acid substitution. The interfacial surface area reported for the physiologically relevant *Hs* Rpn8 homodimer is approximately  $2,300 \text{ \AA}^2$  [176] and for the Rpn11-Rpn8 heterodimer around  $2,000 \text{ \AA}^2$  [160] (Table 4-2). These values predict an equivalently large buried area for the CSN5-CSN6 complex, which within our molecular docking experiments (detailed in section 4.6) all the models obtained, had a buried surface area in the range of  $1,430$  to  $1,695 \text{ \AA}^2$ . It must be noted that conceiving an interfacial variant which has a larger effect in disrupting the interface of the CSN5-CSN6 complex, and which in addition, is not effecting the potential activating property of CSN6 on CSN5 was a great challenge. More specifically, if the mode of interaction for the CSN5-CSN6 heterodimer is as we suggest, the surface largely implicates the Ins-1 segment of CSN5, which from our previous work in Chapter 3, is likely to undergo conformational changes to form an active enzyme. As such targeting further mutations in this region would inevitably have a consequence in the conformation of this segment that is difficult to predict at the atomic level and thus in the activity of CSN5. Nevertheless predicting further conformational changes is also challenging, residues in the interface may be implicated in the interaction and/or in inducing a conformationally active enzyme.

Further to looking into the effect of interfacial mutations on the association of CSN5 and CSN6, we probed their effect on the CSN6-mediated CSN5 activation. Our work has established that *in vitro* the CSN5-CSN6 complex can pursue peptidase activity to process pro-Nedd8 and we have employed this activity as a convenient enzymatic tool.

Pro-Nedd8 processing assays were employed to address the effect of the CSN6 variants in the hydrolytic activity of the CSN5 R106T-CSN6 complex. In addition to the H44A and V115E variants of CSN6, further mutations of residues N85 and E88 to lysine and alanine, respectively, were performed. Both CSN6 N85 and E88 located on the  $\beta$ -sheet  $\beta_3$  are highly conserved residues (Fig. 4-9A) and are on the surface of CSN6 which we propose to mediate the CSN5-CSN6 complex. Our data here show that the hydrolytic activity of the complex is completely suppressed in the presence of the H44A and V115E variants of CSN6 (Fig. 4-9C). Whilst a slight decrease in activity was observed in the case of N85K variant, for the E88A variant, a WT level of activity was observed. These results indicate that, in agreement with the ITC data (ITC experiments were not performed for the N85K and E88A mutants due to the low effect in activity detected), the disruption of the heterodimer interface influences the

catalytic activity. Moreover the strong effect observed from the variants located on helices  $\alpha 1$  and  $\alpha 2$ , compared to those on the  $\beta 3$  strand can suggest that the helices are key in the binding interface of the complex whilst the  $\beta$  strands may be involved in forming hydrophobic patches.

We further utilised pro-Nedd8 processing activity to investigate the contribution of the Ins-1 segment of CSN5 in the interface with CSN6. Mutations were performed on this segment of CSN5 R106T to form the variants Y120D, E122A, Y116D and E115A. Whilst E115A and E122A face the inner catalytic groove, the remaining two variants face the out side (Fig. 4-9D). Pro-Nedd8 processing ability of these variant showed that they were all negatively affected, however to different extents. The activity with CSN5 R106T Y120D and E115A variants resulted in a 30% drop in activity, whilst the E122A and Y116D showed complete inhibition of pro-Nedd8 processing activity (Fig. 4-9E). Although the results here clearly indicate that the Ins-1 region plays a key role in the CSN5-CSN6-Nedd8 ternary complex and we previously showed that it is important to form a catalytically competent CSN5. As such, differentiation between residues implicated in the interaction with CSN6 and those with Nedd8 was proved challenging. However comparison with data presented in section 4.6.1, Fig. 4.10, where we test pro-Nedd8 processing activity of the CSN5 variants in the absence of CSN6, we see that variants, E115A, Y116D and Y120D show approximately 50 % drop in activity, whilst E122A shows WT activity. Together with the results presented in this section, we can conclude that whilst the activity of E115A and Y116D is rescued in the presence of CSN6, the activity of variant Y116D is not and nor can we detect any activity for E122A. The reason we do not detect any activity for E122A is because of concentration differences. Whilst in the absence of CSN6 we use 900  $\mu$ M of enzyme, in the presence of CSN6 we are nanomolar range to obtain 4 nM of CSN5-CSN6 complex. These results thus indicate that residues Y116D and E122A are likely to affect the binding to CSN6 (as CSN6 was also shown not to interact significantly directly with Nedd8).

In this work, we have gathered evidence to propose that the CSN5-CSN6 heterodimer is mediated by a surface delineated by the helices  $\alpha 1$  and  $\alpha 2$  and previously observed in the structures of the Rpn8 homodimer and the Rpn11-Rpn8 heterodimer. We were further able to determine that residues H44 and V115 of CSN6 and residues, Y116 and E122 of CSN5 are likely to be involved in the interaction mediating the CSN5-CSN6 interface. Here we have obtained experimental proof that we further exploited to reconstruct a molecular model of the CSN5-CSN6 complex.

## **4.6 Construction of the ternary CSN5/CSN6/Nedd8 complex:**

The experimental data obtained from characterising the interaction interface for CSN5 with Nedd8 and CSN5 with CSN6 individually prompted us to build a model of this

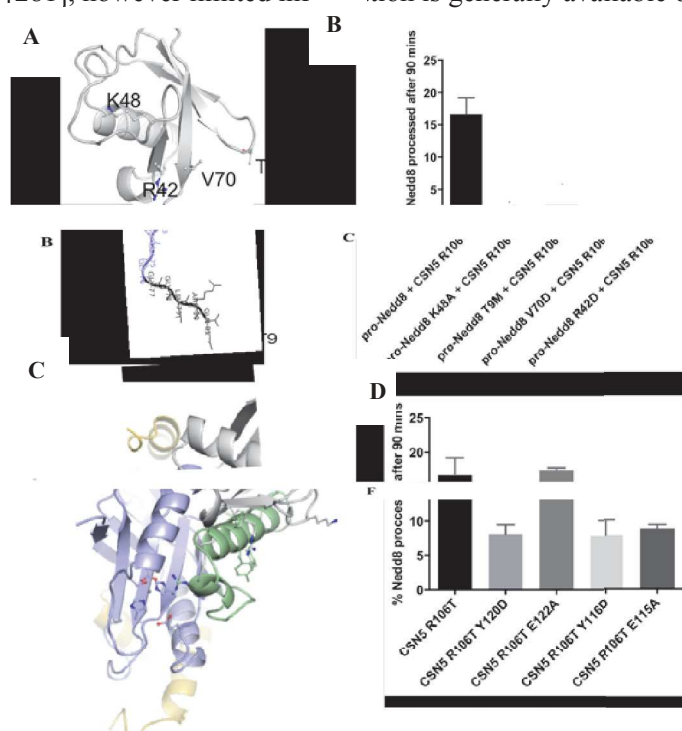
complex using a molecular docking approach with experimental constraints to reconstitute the CSN5-CSN6 binary complex. By performing molecular docking, we aim to predict the three-dimensional structure of the intermolecular complex formed between two or more individual proteins in aqueous solution. Prior to exploring this approach certain challenges and pose limitations of protein-protein docking had to be considered. The difficult cases for docking involve: (i) large-scale conformational movements upon binding; (ii) weak or transient binding; (iii) unavailability of X-ray structure for proteins implicated. Expectedly, such difficulties lead to a considerable drop in the reliability of the predictions obtained.

When employing molecular docking to obtain a putative model corresponding to the CSN5-CSN6 hetero-complex, we thought critically about the highly probable conformational changes occurring upon the interaction between the two protein subunits. In Chapter 3, we showed via MD simulations and structural analysis that the CSN5 subunit contains highly flexible regions, including its Ins-1 and Ins-2 regions. Earlier in this chapter we demonstrated that auto-inhibited CSN5 is activated upon its interaction with CSN6. This finding is suggestive that at least the Ins-1 region of CSN5 would have changed conformation upon the interaction and further conformational changes cannot be excluded. Moreover potential flexibility may also be present in CSN6. The NMR HSQC spectrum of <sup>15</sup>N-uniformly labelled CSN6 indicated conformational changes occurring upon titration of CSN5. Taken together, these results guided us to design an approach where in addition to providing experimental constraints from the experimental methods discussed previously, we strongly consider the conformational changes potentially occurring upon binding. The following sections detail the consecutive steps taken to obtain a molecular model of the CSN5-CSN6 complex.

#### **4.6.1 Investigation of Nedd8 and CSN5 interaction**

The first step in this process was to build a CSN5-Nedd8 complex. To do so, we first determined whether Nedd8 associates with CSN5 in the same way as the distal Ub does with AMSH-LP. The co-crystal structure of AMSH-LP with a di-K63-Ub moiety brought important insights into Ub binding to JAMM/MPN+ proteins (PDB code: 2ZNV). This indicated how specific recognition of Ub and of K63-linked di-ubiquitin is achieved by the direct contact formed between the DUB domain and all the residues involved in the K63 linkage, including K63 itself [186]. This structure allowed the identification of a substrate-binding mode potentially common to other JAMM/MPN+ DUBs. In general, Ub-binding DUBs contain a substrate-binding site that encompasses an extensive interaction network. The foremost interactions are established through the Ub I44 patch, an Ub canonical protein interaction site [264], and additional surfaces specific for this DUB family. Commonly in interactions with DUBs, approximately 20-40% of the total Ub surface can be occupied [217]. The C-terminus of the distal Ub moiety extends from its binding groove and inserts into the

DUB catalytic centre, positioning firmly for catalysis (Chapter 3, Fig. 3-3B). Moreover it is this C-terminal amino acid extension that allows to distinguish between different substrates, including between Ub and Ubl molecules. The C-terminal tail of Ub contains the sequence L71-R72-L73-R74-G75-G76, which differs to that of Nedd8, L-A-L-R-G-G, and to that of SUMO, Q-Q/E-Q-T-G-G. Biochemical studies have demonstrated that the composition of the five last C-terminal residues determines the specificity for the recognition by individual DUB families. For example, AMSH-LP and Rpn11 specifically recognise Ub, whilst CSN5 shows specificity for Nedd8. To date, Nedd8 has only been co-crystallised with NEDP1 [281], however limited information is generally available on Nedd8-bound DUBs.



**Figure 4-10. Nedd8 binding to CSN5.**

**A.** The structure of pro-Nedd8 (PDB code: 2BKR) was extended at the C-terminus to form pro-Nedd8. The model was built in Chimera. Residues that have been mutated in our work have been highlighted and shown in stick mode. **B.** Pro-Nedd8 processing activity assays were employed to characterise the interaction between CSN5 and Nedd8. The Nedd8 variants were tested for pro-Nedd8 processing activity. The plotted bar charts show the percentage of Nedd8 processed after 90 minutes incubation at 37 °C. **C.** Model of CSN5 and Nedd8 interactions. Residues that have been mutated both in pro-Nedd8 and CSN5 R106T are shown in stick mode. **D.** Pro-Nedd8 activity tested with the variants of CSN5 R106T. The Ins-1 located CSN5 R106T variants in complex with CSN6 were tested for pro-Nedd8 processing activity. The plotted bar charts show the percentage of Nedd8 processed after 90 minutes incubation at 37 °C.

We thus sought to characterise the CSN5 mode of Nedd8 binding in more detail. High  $K_D$  values of CSN5 for Nedd8 (as determined by fluorescence anisotropy) deterred us from attempting CSN5/Nedd8 co-crystallisation. Instead, we used mutagenesis

analysis coupled to activity assays. As highly expected, if CSN5 binds to Nedd8 in a similar orientation as does AMSH-LP to Ub, investigation of an AMSH-LP-like binding surface in CSN5 would be expected to have an effect on the binding and potentially in its activity. To test this hypothesis, we generated the variants on pro-Nedd8 including, T9M, R42D, K48D, V70D (Fig. 4-10A and B). The use of pro-Nedd8 here is explained by the fact that it is readily expressed in bacteria (unlike Nedd8 which recombinant expression is notoriously difficult), that it can be easily mutagenised and that we have found that pro-Nedd8 is an *in vitro* substrate for CSN5. These mutations were designed after structural analysis was performed using the structure of di-Ub bound AMSH-LP and their effect was investigated by the pro-Nedd8 processing activity of the active R106T variant of CSN5 (in the absence of CSN6). The rationale behind this approach was also to obtain biochemical evidence that could guide the mapping of Nedd8 binding in the active site of CSN5. The quantification of the bands from our activity assays indicated that all four variants of pro-Nedd8 have decreased catalytic activity. This reduced activity is most evident for the K48D and V70D variants of pro-Nedd8, which completely inhibit the processing activity of CSN5 R106T (Fig. 4-10B). These results confirm our hypothesis that Nedd8 binds to CSN5 in similar manner to Ub binding to AMSH-LP (Fig. 4-10C).

Moreover the CSN5 variants located on the Ins-1 segment, previously designed to detect the implication of this region in the interaction with CSN6, were used to determine whether the Ins-1 region is also involved in the interaction with the substrate. To unravel this, the effect of the following CSN5 R106T variants on the pro-Nedd8 processing activity - E115A, Y116D, Y120D and E122A was tested (Fig 4-9D; 4-10D). The E115A, Y116D, and Y120D variants revealed a significant drop in pro-Nedd8 processing ability compared to CSN5 R106T tested in the same conditions (Fig. 4-10D). However, the CSN5 R106T E122A variant does not affect pro-Nedd8 hydrolysis and this confirmed that E122 does not significantly contribute to substrate binding or catalysis (Fig. 4-10D), but is implicated in the interaction with CSN6 (Fig. 4-9E). Together these results allow to conclude that the Ins-1 segment and more specifically the start of its C-terminal helix and the face locating residues Y116 and Y120 are implicated in the recruitment of substrate or its correct positioning (Fig. 4-9D). The effect observed may be a result of either the inhibition of enzyme-substrate contact or potentially an impact on the flexibility of the Ins-1 segment of CSN5. As we discuss in Chapter 3, the conformation and the dynamic nature of the Ins-1 region is important for the activity of the enzyme.



#### 4.6.2 Model of CSN5-Nedd8 complex

In Chapter 3, we identified that the auto-inhibited CSN5 structure is unable to accommodate Nedd8 accounting for its lack of isopeptidase activity. Detailed structural analysis together with insights from MD simulations allowed us to design the R106T variant of CSN5 exhibiting a conformationally relaxed Ins-1 region. We further showed that this variant is able to bind to Nedd8 and displays significant catalytic activity. Earlier in this chapter we verified our hypothesis that Nedd8 binds to CSN5 in similar manner to Ub binding to AMSH-LP. These findings highlight the challenges that could be faced in using CSN5 for our docking experiments regarding the likelihood that it undergoes conformational changes and most certainly at its Ins-1 region. To aid the docking procedure we thus chose to assemble a complex composed of CSN5 harbouring an Ins-1 conformation similar to that of AMSH-LP, and Nedd8 occupying the same surface to that of AMSH-LP-bound distal Ub. An interesting observation encountered from the modelling of the CSN5 Ins-1 insertion was its remarkable resemblance *a posteriori* to the Ins-1 segment of Rpn11 when in a binary complex with Rpn8.

The generated CSN5-Nedd8 model was the starting complex structure used to perform the molecular docking with CSN6 (described in the next section). This rather provided an extra layer of constraints for the docking approach.

#### 4.6.3 Model of the ternary complex, CSN5-CSN6-Nedd8

In eukaryotes, the heterodimerisation of JAMM/MPN<sup>+</sup> and MPN<sup>-</sup> subunits appears to be an important principle in the regulation of several multiprotein complexes, including the 19S proteasome lid, the BRCA1-RAP80 complex, the BRISC complex and within our work this aspect was demonstrated for the CSN complex. The crystal structures of MPN domain-containing proteins depicted their propensity to form homodimers. As we discussed however, the different MPN-containing domains engage through unique dimerisation interfaces, making the elucidation of which interface corresponds to that observed in heterodimers of JAMM/MPN<sup>+</sup> with MPN<sup>-</sup> subunits challenging to decipher. The elucidation of the human CSN6 structure showing a homodimerisation interface reminiscent to that observed in *HsRpn8* crystal structure hinted a common MPN dimerisation interface. More recently, the structure of the CSN5-CSN6 paralog dimer, Rpn11-Rpn8, from yeast, further depicted an organisation following the mode of dimerisation originally described in the Rpn8 homodimer.

Our initial attempt to investigate the CSN5-CSN6 heterodimerisation by *ab initio* molecular docking using experimental constraints corresponded to the use of the CSN5-Nedd8 model generated and the human CSN6 crystal structure. The molecular

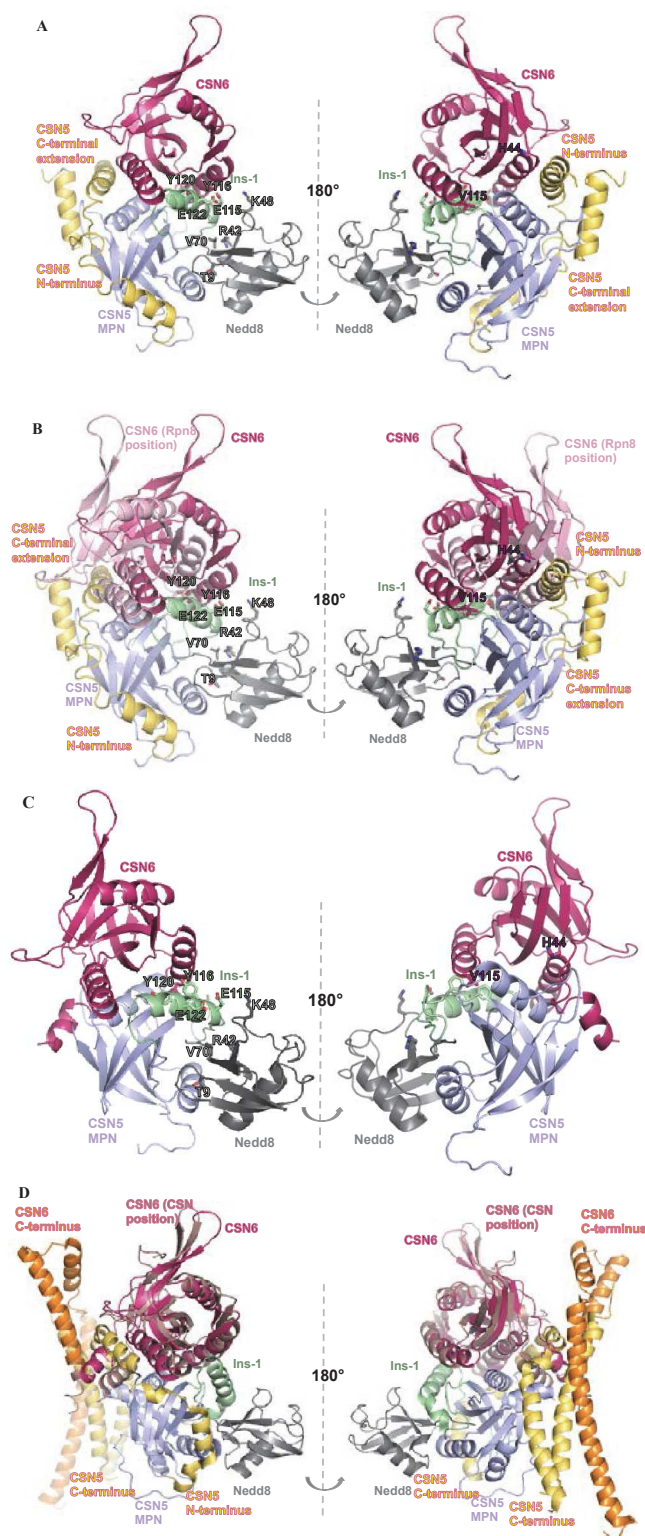


docking was preformed using ZDOCK which utilises a fast Fourier transform to search for all possible binding modes for the proteins. Evaluation of the models generated is based on shape complementarity (defining the match of the convex surface and concave surface of the proteins), desolvation energy, and electrostatics, giving an output of the top 2000 predictions. The predictions of the models generated by ZDOCK were further enhanced by a tailored post-docking scoring strategy resulting in several convincing solutions. The putative solutions were further evaluated by our experimental constraints, resulting in a few assemblies which satisfied data obtained from the NMR and mutagenesis analysis performed previously (Fig. 4-11A).

Interestingly, comparison between the crystal structure of the Rpn11-Rpn8 complex and the selected docking complex for CSN5-CSN6 revealed to be highly similar. Although occupying the same protein-protein interaction surface the main difference was depicted in the topology of the CSN6 subunit in comparison to that of Rpn8. Relative to the position of Rpn8 on Rpn11, the CSN6 subunit in our selected model shows a rotation of 25° and is translated relative to the Rpn8 position (Fig 4-11B). This observation indicates a deviation from the two-fold symmetry observed in the structure of Rpn11-Rpn8, but also of the Rpn8 and human CSN6 homodimers. Structural analysis identified that the cause for this difference in the CSN6 topology accounted to the presence of the N- and C-terminal helical appendices corresponding to residues, 6 to 71 and 235 to 251, respectively, in CSN5. As previously mentioned in Chapter 3 (section 3.2.1.2, Fig. 3-4A) in the CSN5 structure these two helices are stapled together via a hydrophobic core. We had investigated the flexibility of the N- and C-terminal extensions of CSN5 by MD simulations and had identified these regions as highly flexible (referring to Chapter 3, Fig. 3-8C). This data was suggestive of potential conformational changes occurring at this region of the N- and C-terminal helices. Additionally compared to the N-terminal helix of Rpn11, this helix is slightly longer in CSN5 and in particular carries the characteristic of wrapping around the MPN core of CSN5. Taken together we thus hypothesised that the reason for the deviation observed between the structural topology of the Rpn11-Rpn8 and that of the modelled CSN5-CSN6 complex is likely owed to the N- and C-terminal appendices of CSN5 preventing CSN6 to be placed in the exact same position as Rpn8 on Rpn11.

We furthered our molecular docking simulations to investigate the importance of the N- and C-terminal regions of CSN5 in the assembly of the heterodimer complex (Fig 4-11C). We thus repeated the procedure with ZDOCK using a CSN5 variant lacking the N- and C-terminal helices. Following our post-docking scoring strategy and further validation with our experimental constraints, the top scoring solution generated a model of the CSN5-CSN6 complex strongly superimposable to that of the Rpn11-Rpn8 complex. An rmsd value of 2.5 Å was obtained for these two complexes. Most interestingly, upon the report of the CSN structure, superimposition of our putative CSN5-CSN6 model with the CSN5-CSN6 structure of the CSN (Fig 4-11D),

identified highly coherent fit, giving an rmsd value of 1.8 Å for the superimposition of the CSN6 subunits within their heterodimer complex.



**Figure 4-11. JAMM/MPN+-MPN- heterodimer assembly.**

Residues discussed in the text are shown in stick mode. **A.** The best docking pose model of CSN5-Nedd8-CSN6 obtained by data-driven docking. Right panel: 180° rotated view of left panel. **B.** Topology of CSN5-CSN6 docking model and Rpn11-Rpn8-like CSN5-CSN6. The best CSN5-Nedd8-CSN6 pose obtained from filtering ZDOCK solutions is presented here. To this model, the CSN5-Nedd8-CSN6 model built following the Rpn11-Rpn8 (PDB code: 4O8X) structure has been superimposed. The CSN6 structure superimposed on Rpn8 is shown in light pink to distinguish it from the position in our docking solution. Right panel: 180° rotated view of left panel. **C.** Removal of N- and C-terminal extensions and the resulting docking pose. The best docking pose using a CSN5<sup>ΔC</sup> model in which the N- and C-terminal extensions were deleted is shown. Right panel: 180° rotated view of left panel. **D.** Comparison with the CSN5-CSN6 heterodimer in the CSN complex. The docking pose obtained for the CSN5-CSN6 heterodimer using the CSN5<sup>ΔC</sup> model is superimposed with the CSN5-CSN6 from the crystal structure of the CSN complex (PDB code: 4D18). Right panel: 180° rotated view of left panel.

Here we have constructed a putative model of the CSN5-CSN6 heterodimer showing a protein-protein interaction surface that is also found in the Rpn11-Rpn8, Rpn8, and human CSN6 dimers, and is also in good agreement with the CSN5-CSN6 crystal structure from the CSN complex recently published. Our findings emphasise the determinant role of the position of the N- and C-terminal helices of CSN5, where their dissociation is required for CSN5 and CSN6 to interact through an architecture reminiscent of Rpn11-Rpn8. Following our hypothesis that Nedd8 interacts with CSN5 in a similar manner as distal Ub associates with AMSH-LP, we constructed a putative model of the CSN5-CSN6-Nedd8 ternary complex.

## Chapter 5: Conclusions and Discussion

In the light of our work, we have compared the CSN complex and, more specifically, its catalytic subunit, CSN5, with various model systems. The crystal structure of the inactive CSN5 (fragment 1-257) subunit was initially solved by the group and, at the time, was the only JAMM/MPN<sup>+</sup> protein crystallised from the ‘Zomes’ family. Our work initially encompassed the biophysical characterisation of the catalytic subunit (Fig. 5-1A). Sequence alignment of CSN5 from different organisms reveals highly conserved features, identifying key components of its catalytic function within the CSN complex. A number of studies have proposed the presence of CSN5 distributed between the CSN complex, different CSN sub-complexes and in a CSN-independent form. Although this distribution may contribute to a major regulatory element, little is known about the function and regulation of these CSN5-containing complexes – even their physiological significance requires firmer investigations. Clearly for a number of years now, the best understood function of CSN5 and thus of the CSN complex resides in the deneddylation of Cullins. This is this catalytic aspect that we mainly focused on.

Besides its catalytic function within the CSN complex, the CSN5 subunit was suggested to mediate other biochemical purposes, which have not been clearly comprehended. Suggestions of roles of CSN5 include, specificity determination of transcription factors (such as AP-1 (for Activator Protein-1) and E2F-1) [123, 292], regulation of intracellular distribution and turn-over of signalling molecules (p27, COP1, and p53) [107, 293], and promoting c-Jun, NFκB, and p53 phosphorylation [88, 119]. Our work has identified that *in vitro* CSN5 can be found in different oligomeric assemblies and that, in cells, CSN5 is found to be in equilibrium between monomeric and dimeric species. Although CSN5 has been found to interact with various partners as mentioned earlier, the studies that have investigated these interactions have relied solely on a monomeric form of CSN5 being implicated and should be revisited taking into account the ability of CSN5 to auto-associate. Indeed, in our work, our findings led us to propose that the dimeric assembly of this subunit could be relevant in mediating protein-protein interactions. We have further utilised various approaches, ranging from cellular, solution, and *in silico* techniques, to determine structural information regarding the dimeric arrangements and we were able to propose the most probable, physiologically relevant, dimeric interaction interface. Contrary to the homodimeric assembly of CSN6 observed in the crystal, this homodimeric CSN5 interface does not mimic the CSN5-CSN6 heterodimer.

Early on in this work, we have mainly presented our findings on CSN5 activity, mainly focusing on its activation state in the CSN-independent state. The complementary use of structural biology and MD simulations allowed the detailed characterisation of CSN5 activity regulation at the molecular level. Our work

encompassed a detailed comparison of the inactive CSN5 with the catalytically active structure of AMSH-LP/K63-Ub2, where distinctively the catalytic zinc-binding site of both proteins adopts a common geometry. This observation had suggested that, as AMSH-LP, CSN5 active site is competent to perform isopeptidase activity and to hydrolyse an isopeptide bond. CSN5 crystal structure and the investigation of the recruitment of Nedd8 substrate by CSN5 identified that, unlike the AMSH-LP, part of substrate-binding site of CSN5 delineated by the Ins-1 segment would not be sterically able to accommodate its substrate, in particular the LRGG C-terminus segment. This observation led us to suggest that this Ins-1 insertion of the CSN5 would require substantial conformational rearrangement to be primed to recruit Nedd8. This hypothesis is based on the ability of the Ins-1 segment to display a large degree of flexibility. The use of MD simulations allowed to investigate the plasticity of this segment and identified a functional role of the conserved R106 residue. We identified that the R106 residue is implicated in forming a salt bridge with the zinc-chelating D151 residue and by mutational studies, both *in silico* and *in vitro*, we revealed that the breakage of the salt bridge is required for the rearrangement of the Ins-1 region and for restoring some 'basal' isopeptidase activity on the CSN5 subunit. The use of two synthetic substrates, LRGG-AMC and Nedd8-AMC, confirmed that the substitution of R106 to a threonine and, to a lesser extent, to an alanine and glycine (but not proline) imposes constitutive activity to CSN5 in the stand-alone form. Together these results allowed to identify that the strictly conserved R106 residue is a key component in the activation switch of the isolated form of CSN5 (Fig. 5-1C). Comparison between the binding affinity of wild type CSN5 and CSN5 R106T to Nedd8 showed that the increase in flexibility of the Ins-1 insertion segment is indeed associated with a significant enhancement of the binding affinity for the Nedd8 substrate and thus contributes to form a catalytically more efficient enzyme. Furthermore, our work also identified that the R106T variant is also able to deneddylate neddylated Cul1, although at a modest level compared to the CSN activity. However, these results confirmed those obtained using synthetic substrates.

Our work and those of others have suggested that CSN5 activation is highly regulated, keeping CSN5 under tight control. We have shown that CSN5 activation requires an activation step and one such step corresponds to the 'unhooking' of the R106 residue from the zinc-binding site, which in turn promotes the opened conformation of the enzyme and reveals its isopeptidase activity. The integration of CSN5 into the CSN complex and the consequent protein-protein interactions with other CSN subunits were speculated from various studies to play a role in both CSN5 activation as well as substrate recruitment. Most recently, Lingaraju et al. characterised the activation of CSN5 within the crystal structure of the CSN complex. However, it was indeed a surprise that CSN5 was found in an inactive conformation in the CSN crystal. Interestingly whereas the histidine and aspartate residues in the active site of CSN5 are positioned as expected, the catalytic water molecule is replaced by another amino acid, E104. The role of this residue in CSN regulation was tested and findings indicated that E104-Zn binding might serve to keep the CSN in an inactive state when



it is not conjugating to CRLs. These findings together with our work suggest that CSN5 is inhibited by distinct mechanisms depending on whether the subunit is on its own or integrated into the CSN.

Given the evidence from studies on the CSN complex, but also from JAMM/MPN+ subunit-containing complexes, such as the 26S proteasome lid and the BRCC36-containing complexes, we furthered our investigation of CSN5 activation by probing its interaction with the MPN- subunit, CSN6. Although previously highly speculated, we provided supporting evidence that the two subunits interact directly via their MPN domains. We further structurally, biophysically and functionally characterised the interaction *in vitro*. Upon near completion of this work two recent reports on the paralog MPN+-MPN- heterodimer of Rpn11-Rpn8 were conveyed with data on the crystal structure of the complex and on its activity. The design of these studies prevented the quantification of the interaction affinity of the MPN domain proteins, as they were either co-expressed or expressed as a single polypeptide chain, because of the poor stability of Rpn11 when expressed alone. In our work, we were therefore able to go further and quantify the interaction of CSN5 and CSN6 MPN domains to a  $K_D$  of  $\sim 6 \mu\text{M}$ , although the C-terminal extensions of the two proteins, absent in our constructs, could play a role in the subunit interaction affinity. The association of CSN6 to CSN5 was shown to markedly enhance its hydrolytic activity on a variety of substrates and to increase the affinity for Nedd8 (Fig. 5-1D). Quantification of the kinetic parameters describing the activity of the CSN5-CSN6 complex and comparison with other isopeptidases, including the CSN complex suggest that the MPN heterodimer could contribute significantly to the catalytic activity of the human CSN, potentially forming the core catalytic domain of this complex. Although, for Rpn11 and BRCC36, it had previously been shown that their isopeptidase activity requires with the presence of their corresponding MPN- domain-containing subunits [168], contradictory data for CSN5 was reported [143]. In this work, the function of CSN6/Csi1 in human and in budding yeast, respectively was probed and conclusions derived suggested that only the C-terminal region of CSN6 is necessary for the deneddylation activity. Although highly confident in our results obtained as various different substrates were used to detect the activity in the presence of CSN6, further investigation may be required to reconcile the different results obtained in these two studies.

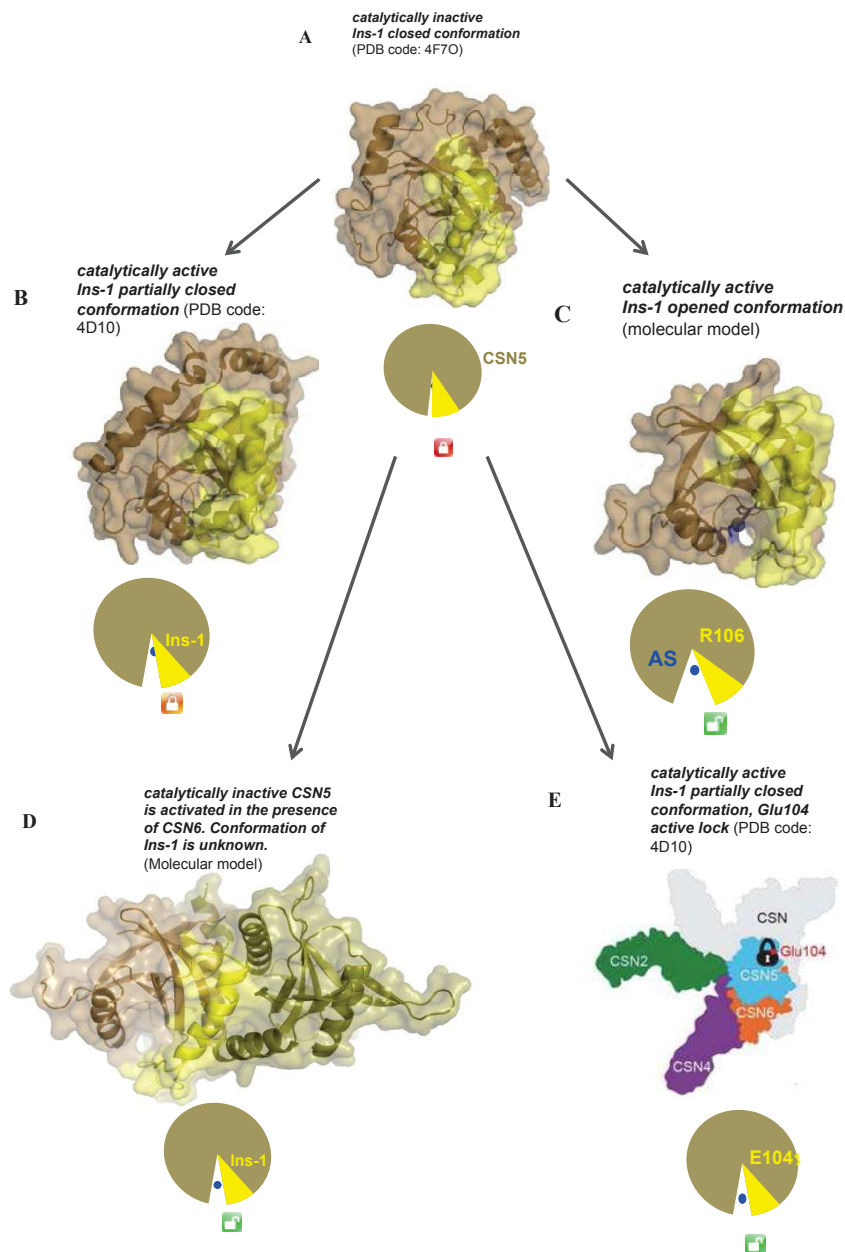
Although attempts to characterise the CSN5-CSN6 interaction by X-ray crystallography were undertaken within our work as well, whilst this approach claimed unsuccessful, we also investigated the interaction by means of an integrated approach comprising of interface mapping by NMR, sequence conservation analysis, site-directed mutagenesis combined with enzymatic activity assays. These experimental approaches provided experimental constraints which were further used to reconstitute the CSN5-CSN6 heterodimer complex by molecular docking. Interestingly, the surfaces engaged in the CSN5-CSN6 model produced are highly similar to those observed in the Rpn11-Rpn8 complex although differences in the C-

terminal region were detected. This probed us to identify and characterise the flexibility of the N- and C-terminal regions of CSN5. We thereby showed that the N- and C-terminal regions of CSN5, adopting a stapled conformation in the crystal structure, has the ability to undergo conformational changes. In that case, the resulting CSN5-CSN6 heterodimer model obtained is reminiscent to that of the yeast Rpn11-Rpn8 heterodimer, and human Rpn8 and CSN6 homodimers mode of interaction.

Our work provided an updated model of the CSN5-CSN6 complex and overall highlighted the importance of the MPN domain interaction in the catalytic activity. This could further suggest that the C-terminal regions of full length CSN5 and CSN6 may contribute to anchoring and optimal positioning of these subunits within the CSN complex. Whilst we were in the processing of submitting our work, a manuscript reporting the crystal structure of human CSN complex at 3.8 Å resolution was published [61]. Findings described in this study are in good agreement and support conclusions reached in our study.

Within the thesis occasionally we have briefly mentioned and compared the work presented in this study to ours. However detailed investigation in this manuscript was limited as a result of the timing of this paper barely coincided with our work. Nevertheless this study, together with others which we have detailed earlier in our work, and of course with our present studies, have contributed in the blossom of information regarding the activity and activity regulation of this protein complex. The structure of the CSN complex depicted some anticipated features coming from previous work on related proteins. However the inactive configuration of the CSN in the crystal structure was a surprise for many (Fig. 5-1B). A further difference was observed in the Zn coordination site, where the activated water molecule is replaced by another amino acid, E104 from the Ins-1 insertion segment. Interestingly, Lingaraju *et al.* proposed that the E104-Zn<sup>2+</sup> binding might be responsible for keeping the CSN in an inactive state when free from CRLs (Fig. 5-1E). These findings together with ours, would suggest that the CSN5 subunit is inhibited by distinct mechanisms depending on whether the subunit is in its stand-alone state or whether within the CSN complex. These findings reveal that, although some generalisations could be made across the JAMM/MPN+ family, they all seem to have unique and distinctive features. These findings not only contribute to answering a number of questions which have been awaiting for many years, but also open further doors in the field of JAMM/MPN+ family proteins, and much more exciting work remains to be done to grasp the intricate regulatory mechanisms responsible for the controlled functions of the CRLs and more globally, of the UPS.





**Figure 5-1. Modes of activity regulation of CSN5 and the CSN complex.**

**A.** CSN5 alone is inactive, the Ins-1 segment is in a closed conformation, blocking the site for Nedd8 binding. **B.** Ins-1 of CSN5 within the CSN complex, is in a partially closed conformation. **C.** A molecular model of CSN5 where the Ins-1 is in an open conformation, due to breaking of R106-D151 salt bridge, allowing for substrate accommodation. **D.** CSN5 is activated by CSN6, the Ins-1 is speculated to be in an open conformation. **E.** The E104-Zn<sup>2+</sup> binding is responsible for keeping the CSN in an inactive state when free from CRLs. Upon CRL binding, the CSN is activated.

Even from the structural point of view, there still remains work ahead including for example the examination of the structure of different functional mutants of the CSN, which could help unravel the mechanism by which binding of the CRLs induces

conformational changes. Even more intriguing would be to investigate the CSN-Nedd8-CRL complex and determine the CSN-CRL interface and the influence of Nedd8 in the activation of the CSN. Moreover, the structural studies of CSN5, CSN5-CSN6 and the CSN complex, now provide templates for the design of drugs that act on the CSN, which could be an attractive target for various cancers.

## Bibliography

1. Koepp, M.D., W.J. Harper, and J.S. Elledge, *How the Cyclin Became a Cyclin: Regulated Proteolysis in the Cell Cycle*. Cell, 1999. **97**: p. 431-434.
2. Joazeiro, P.A.C. and M.A. Weissman, *RING Finger Proteins: Mediators of Ubiquitin Ligase Activity*. Cell, 2000. **102**: p. 549-552.
3. Hicke, L., *Protein Regulation by Monoubiquitin*. Molecular Cell Biology, 2001. **2**: p. 195-201
4. Rock, L.K. and L.A. Goldberg, *Degradation of cell proteins and the generation of MHC class I-presented peptides*. Annu. Rev. Immunol., 1999. **17**: p. 739-779.
5. Hershko, A. and A. Ciechanover, *The ubiquitin system*. Annu. Rev. Biochem., 1998. **67**: p. 425-479.
6. Hochstrasser, M., *Evolution and function of ubiquitin-like protein-conjugation systems*. Nature Cell Biology 2000. **2**: p. 153-157.
7. Pickart, M.C., *Mechanisms underlying ubiquitination*. Annu Rev Biochem, 2001. **70**: p. 503-533.
8. Thrower, S.J., et al., *Recognition of the polyubiquitin proteolytic signal*. The EMBO Journal, 2000. **19**: p. 94-102.
9. Chan, L.N. and P.C. Hill, *Defining polyubiquitin chain topology*. Nature Structural Biology 2001. **8**: p. 650-652.
10. Haglund, K., P.P. Di Fiore, and I. Dikic, *Distinct monoubiquitin signals in receptor endocytosis*. Trends Biochem Sci, 2003. **28**(11): p. 598-603.
11. Hofmann, M.R. and M.C. Pickart, *Noncanonical MMS2-Encoded Ubiquitin-Conjugating Enzyme Functions in Assembly of Novel Polyubiquitin Chains for DNA Repair*. Cell, 1999. **96**: p. 645-653.
12. Deng, L., et al., *Activation of the I $\kappa$ B Kinase Complex by TRAF6 Requires a Dimeric Ubiquitin-Conjugating Enzyme Complex and a Unique Polyubiquitin Chain*. Cell, 2000. **103**: p. 351-361.
13. Spence, J., et al., *Cell Cycle-Regulated Modification of the Ribosome by a Variant Multiubiquitin Chain*. Cell, 2000. **102**: p. 67-76.
14. Tokunaga, F., et al., *Involvement of linear polyubiquitylation of NEMO in NF-kappaB activation*. Nat Cell Biol, 2009. **11**(2): p. 123-32.
15. Scaglione, K.M., et al., *The ubiquitin-conjugating enzyme (E2) Ube2w ubiquitinates the N terminus of substrates*. J Biol Chem, 2013. **288**(26): p. 18784-8.
16. Tatham, M.H., et al., *Ube2W conjugates ubiquitin to alpha-amino groups of protein N-termini*. Biochem J, 2013. **453**(1): p. 137-45.
17. Adams, J., et al., *Proteasome Inhibitors: A Novel Class of Potent and Effective Antitumor Agents*. Cancer Res, 1999. **59**: p. 2615-2622.
18. Deshaies, R.J. and C.A. Joazeiro, *RING domain E3 ubiquitin ligases*. Annu Rev Biochem, 2009. **78**: p. 399-434.
19. Ozkan, E., H. Yu, and J. Deisenhofer, *Mechanistic insight into the allosteric activation of a ubiquitin-conjugating enzyme by RING-type ubiquitin ligases*. Proc Natl Acad Sci U S A, 2005. **102**(52): p. 18890-5.
20. Li, W., et al., *Genome-Wide and Functional Annotation of Human E3 Ubiquitin Ligases Identifies MULAN, a Mitochondrial E3 that Regulates the Organelle's Dynamics and Signaling*. PLOS ONE, 2008. **3**: p. e1487.

21. Huibregtse, M.J., et al., *A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase*. Proc. Natl. Acad. Sci. USA, 1995. **92**: p. 2563-2567.
22. Wenzel, D.M. and R.E. Klevit, *Following Ariadne's thread: a new perspective on RBR ubiquitin ligases*. BMC Biol, 2012. **10**: p. 24.
23. Budhidarmo, R., Y. Nakatani, and C.L. Day, *RINGs hold the key to ubiquitin transfer*. Trends Biochem Sci, 2012. **37**(2): p. 58-65.
24. Petroski, M.D. and R.J. Deshaies, *Function and regulation of cullin-RING ubiquitin ligases*. Nat Rev Mol Cell Biol, 2005. **6**(1): p. 9-20.
25. Kosarev, P., K.F. Mayer, and C.S. Hardtke, *Evaluation and classification of RING-finger domains encoded by the Arabidopsis genome*. Genome Biology, 2002. **3**: p. 1-12.
26. Lydeard, J.R., B.A. Schulman, and J.W. Harper, *Building and remodelling Cullin-RING E3 ubiquitin ligases*. EMBO Rep, 2013. **14**(12): p. 1050-61.
27. Skaar, J.R., J.K. Pagan, and M. Pagano, *Mechanisms and function of substrate recruitment by F-box proteins*. Nat Rev Mol Cell Biol, 2013. **14**(6): p. 369-81.
28. Soucy, T.A., et al., *An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer*. Nature, 2009. **458**(7239): p. 732-6.
29. Cardozo, T. and M. Pagano, *The SCF ubiquitin ligase: insights into a molecular machine*. Nat Rev Mol Cell Biol, 2004. **5**(9): p. 739-51.
30. Jin, J., et al., *Systematic analysis and nomenclature of mammalian F-box proteins*. Genes Dev, 2004. **18**(21): p. 2573-80.
31. Wu, K., et al., *The SCFHOS/ $\beta$ -TRCP-ROC1 E3 Ubiquitin Ligase Utilizes Two Distinct Domains within CUL1 for Substrate Targeting and Ubiquitin Ligation*. Mol. Cell. Biol., 2000. **20**: p. 1382-1393.
32. Nakayama, K.I. and K. Nakayama, *Ubiquitin ligases: cell-cycle control and cancer*. Nat Rev Cancer, 2006. **6**(5): p. 369-81.
33. Bosu, D.R. and E.T. Kipreos, *Cullin-RING ubiquitin ligases: global regulation and activation cycles*. Cell Div, 2008. **3**: p. 7.
34. Osaka, F., et al., *A new NEDD8-ligating system for cullin-4A*. Genes & Development, 1998. **12**(15): p. 2263-2268.
35. Gong, L. and E.T.H. Yeh, *Identification of the Activating and Conjugating Enzymes of the NEDD8 Conjugation Pathway*. Journal of Biological Chemistry, 1999. **274**(17): p. 12036-12042.
36. Zheng, J., et al., *CAND1 Binds to Unneddylated CUL1 and Regulates the Formation of SCF Ubiquitin E3 Ligase Complex*. Molecular Cell 2002. **10**: p. 1519-1526.
37. Sakata, E., et al., *Direct interactions between NEDD8 and ubiquitin E2 conjugating enzymes upregulate cullin-based E3 ligase activity*. Nat Struct Mol Biol, 2007. **14**(2): p. 167-8.
38. Wu, K., et al., *DEN1 is a dual function protease capable of processing the C terminus of Nedd8 and deconjugating hyper-neddylated CUL1*. J Biol Chem, 2003. **278**(31): p. 28882-91.
39. Kawakami, T., et al., *NEDD8 recruits E2-ubiquitin to SCF E3 ligase*. The EMBO Journal, 2001. **20**: p. 4003-4012.
40. den Besten, W., et al., *NEDD8 links cullin-RING ubiquitin ligase function to the p97 pathway*. Nat Struct Mol Biol, 2012. **19**(5): p. 511-6, S1.

41. Bandau, S., et al., *UBXN7 docks on neddylated cullin complexes using its UIM motif and causes HIF1alpha accumulation*. BMC Biol, 2012. **10**: p. 36.
42. Liu, J., et al., *NEDD8 Modification of CUL1 Dissociates p120CAND1, an Inhibitor of CUL1-SKP1 Binding and SCF Ligases*. Molecular Cell, 2002. **10**: p. 1511–1518.
43. Min, K.W., et al., *TIP120A associates with cullins and modulates ubiquitin ligase activity*. J Biol Chem, 2003. **278**(18): p. 15905-10.
44. Goldenberg, S.J., et al., *Structure of the Cand1-Cul1-Roc1 complex reveals regulatory mechanisms for the assembly of the multisubunit cullin-dependent ubiquitin ligases*. Cell, 2004. **119**(4): p. 517-28.
45. Schmidt, M.W., et al., *F-box-directed CRL complex assembly and regulation by the CSN and CAND1*. Mol Cell, 2009. **35**(5): p. 586-97.
46. Enchev, R.I., et al., *Structural basis for a reciprocal regulation between SCF and CSN*. Cell Rep, 2012. **2**(3): p. 616-27.
47. Lyapina, S., et al., *Promotion of NEDD-CUL1 conjugate cleavage by COP9 signalosome*. Science, 2001. **292**(5520): p. 1382-5.
48. Schwechheimer, C. and W.X. Deng, *COP9 signalosome revisited: a novel mediator of protein degradation*. Trends Cell Biol, 2001. **11**: p. 420-426.
49. Groll, M., et al., *Structure of 20S proteasome from yeast at 2.4Å resolution*. Nature, 1997. **386**: p. 463-471.
50. Unno, M., et al., *The Structure of the Mammalian 20S Proteasome at 2.75 Å Resolution*. Structure, 2002. **10**: p. 609–618.
51. Glickman, H.M., et al., *A Subcomplex of the Proteasome Regulatory Particle Required for Ubiquitin-Conjugate Degradation and Related to the COP9 Signalosome and eIF3*. Cell, 1998. **94**: p. 615-623.
52. Frickel, E.M., et al., *Apicomplexan UCHL3 retains dual specificity for ubiquitin and Nedd8 throughout evolution*. Cell Microbiol, 2007. **9**(6): p. 1601-10.
53. Shen, L., et al., *Structural basis of NEDD8 ubiquitin discrimination by the deNEDDylating enzyme NEDP1*. The EMBO Journal, 2005. **24**: p. 1341–1351.
54. Wei, N. and W.X. Deng, *Making sense of the COP9 signalosome a regulatory protein complex conserved from Arabidopsis to human*. Trends Genet., 1999. **15**: p. 98–103.
55. Dessau, M., et al., *The Arabidopsis COP9 signalosome subunit 7 is a model PCI domain protein with subdomains involved in COP9 signalosome assembly*. Plant Cell, 2008. **20**(10): p. 2815-34.
56. Ellisdon, A.M. and M. Stewart, *Structural biology of the PCI-protein fold*. Bioarchitecture, 2012. **2**(4): p. 118-123.
57. Hofmann, K. and P. Bucher, *The PCI domain: a common theme in three multiprotein complexes*. Trends Biochem. Sci., 1998. **23**: p. 204-205.
58. Lee, H.J., et al., *Crystal structure and versatile functional roles of the COP9 signalosome subunit 1*. Proc Natl Acad Sci U S A, 2013. **110**: p. 11845–11850.
59. Estrin, E., et al., *Formation of an intricate helical bundle dictates the assembly of the 26S proteasome lid*. Structure, 2013. **21**(9): p. 1624-35.
60. Forster, F., et al., *Unveiling the long-held secrets of the 26S proteasome*. Structure, 2013. **21**(9): p. 1551-62.

61. Lingaraju, G.M., et al., *Crystal structure of the human COP9 signalosome*. Nature, 2014. **512**(7513): p. 161-5.
62. Tsuge, T., M. Matsui, and N. Wei, *The subunit 1 of the COP9 signalosome suppresses gene expression through its N-terminal domain and incorporates into the complex through the PCI domain*. J Mol Biol, 2001. **305**(1): p. 1-9.
63. Echaliier, A., et al., *Insights into the regulation of the human COP9 signalosome catalytic subunit, CSN5/Jab1*. Proceedings of the National Academy of Sciences, 2013. **110**(4): p. 1273-1278.
64. Zhang, H., et al., *The crystal structure of the MPN domain from the COP9 signalosome subunit CSN6*. FEBS Lett, 2012. **586**(8): p. 1147-53.
65. Ambroggio, X.I., D.C. Rees, and R.J. Deshaies, *JAMM: a metalloprotease-like zinc site in the proteasome and signalosome*. PLoS Biol, 2004. **2**(1): p. E2.
66. Tran, H.J.T.T.A., M. D.; Löwe, J.; Bycroft, M, *Structure of the Jab1/MPN Domain and Its Implications for Proteasome Function*. Biochemistry, 2003. **42**(39): p. 11460-11465.
67. Birol, M.E., Aude, *Structure and Function of MPN (Mpr1/Pad1 N-terminal) Domain- Containing Proteins*. Curr Protein Pept Sci., 2014. **15**.
68. Wang, X., *The COP9 Signalosome Interacts with SCF<sup>UFO</sup> and Participates in Arabidopsis Flower Development*. The Plant Cell Online, 2003. **15**(5): p. 1071-1082.
69. Cope, G.A. and R.J. Deshaies, *Targeted silencing of Jab1/Csn5 in human cells downregulates SCF activity through reduction of F-box protein levels*. BMC Biochem, 2006. **7**: p. 1.
70. Cope, G.A., et al., *Role of predicted metalloprotease motif of Jab1/Csn5 in cleavage of Nedd8 from Cul1*. Science, 2002. **298**(5593): p. 608-11.
71. Deshaies, J.R., D.E. Emberley, and A. Saha, *Control of Cullin - Ring Ubiquitin Ligase Activity by Nedd8*, in *Conjugation and Deconjugation of Ubiquitin Family Modifiers*, G. M., Editor. 2010, Landes Bioscience, Austin, TX. p. 41 - 56.
72. Emberley, E.D., R. Mosadeghi, and R.J. Deshaies, *Deconjugation of Nedd8 from Cul1 is directly regulated by Skp1-F-box and substrate, and the COP9 signalosome inhibits deneddylated SCF by a noncatalytic mechanism*. J Biol Chem, 2012. **287**(35): p. 29679-89.
73. Bornstein, G., D. Ganoth, and A. Hershko, *Regulation of neddylation and deneddylation of cullin1 in SCFSkp2 ubiquitin ligase by F-box protein and substrate*. Proc Natl Acad Sci U S A, 2006. **103**(31): p. 11515-20.
74. Chew, E.H. and T. Hagen, *Substrate-mediated regulation of cullin neddylation*. J Biol Chem, 2007. **282**(23): p. 17032-40.
75. Cope, A.G. and J.R. Deshaies, *COP9 Signalosome: A Multifunctional Regulator of SCF and Other Cullin-Based Ubiquitin Ligases*. Cell, 2003. **114**: p. 663-671.
76. Miyauchi, Y., et al., *The COP9/signalosome increases the efficiency of von Hippel-Lindau protein ubiquitin ligase-mediated hypoxia-inducible factor-alpha ubiquitination*. J Biol Chem, 2008. **283**(24): p. 16622-31.
77. Choo, Y.Y., et al., *Characterization of the role of COP9 signalosome in regulating cullin E3 ubiquitin ligase activity*. Mol Biol Cell, 2011. **22**(24): p. 4706-15.
78. Lee, W.E., W. Oh, and J. Song, *Jab1 as a Mediator of Nuclear Export and Cytoplasmic Degradation of p53*. Mol. Cells, 2006. **22**: p. 133-140.



79. Richardson, K.S. and W. Zundel, *The emerging role of the COP9 signalosome in cancer*. Mol Cancer Res, 2005. **3**(12): p. 645-53.
80. Lee, M.-H., et al., *Roles of COP9 signalosome in cancer*. Cell Cycle, 2011. **10**(18): p. 3057-3066.
81. Adler, A.S., et al., *CSN5 isopeptidase activity links COP9 signalosome activation to breast cancer progression*. Cancer Res, 2008. **68**(2): p. 506-15.
82. Adler, A.S., et al., *Genetic regulators of large-scale transcriptional signatures in cancer*. Nat Genet, 2006. **38**(4): p. 421-30.
83. Shackelford, T.J. and F.X. Claret, *JAB1/CSN5: a new player in cell cycle control and cancer*. Cell Div, 2010. **5**: p. 26.
84. Zhao, R., et al., *Subunit 6 of the COP9 signalosome promotes tumorigenesis in mice through stabilization of MDM2 and is upregulated in human cancers*. J Clin Invest, 2011. **121**(3): p. 851-65.
85. Lee, Y.H., et al., *Molecular targeting of CSN5 in human hepatocellular carcinoma: a mechanism of therapeutic response*. Oncogene, 2011. **30**(40): p. 4175-84.
86. Kiichiro Tomoda, Yukiko Kubota, and J.-y. Kato, *Degradation of the cyclin-dependent-kinase inhibitor p27Kip1 is instigated by Jab1*. Nature, 1999. **398**: p. 160-165.
87. Yang, X., et al., *The COP9 Signalosome Inhibits p27kip1 Degradation and Impedes G1-S Phase Progression via Deneddylation of SCF Cul1*. Current Biology, 2002. **12**: p. 667-672.
88. Bech-Otschir, D., et al., *COP9 signalosome-specific phosphorylation targets p53 to degradation by the ubiquitin system*. The EMBO Journal, 2001. **20**: p. 1630-1699.
89. Zhang, X.C., et al., *Roles for CSN5 in control of p53/MDM2 activities*. J Cell Biochem, 2008. **103**(4): p. 1219-30.
90. Kim, B.C., et al., *Jab1/CSN5, a Component of the COP9 Signalosome, Regulates Transforming Growth Factor Signaling by Binding to Smad7 and Promoting Its Degradation*. Molecular and Cellular Biology, 2004. **24**(6): p. 2251-2262.
91. Kim, J.H., et al., *Jab1/CSN5 induces the cytoplasmic localization and degradation of RUNX3*. J Cell Biochem, 2009. **107**(3): p. 557-65.
92. Berse, M., et al., *Ubiquitin-dependent degradation of Id1 and Id3 is mediated by the COP9 signalosome*. J Mol Biol, 2004. **343**(2): p. 361-70.
93. Denti, S., et al., *The COP9 signalosome regulates Skp2 levels and proliferation of human cells*. J Biol Chem, 2006. **281**(43): p. 32188-96.
94. Bemis, L., et al., *Distinct aerobic and hypoxic mechanisms of HIF- $\alpha$  regulation by CSN5*. Genes Dev, 2004. **18**(7): p. 739-44.
95. S. Mahalingam, et al., *HIV-1 Vpr interacts with a human 34-kDa mov34 homologue, a cellular factor linked to the G2/M phase transition of the mammalian cell cycle*. PNAS, 1998. **95**: p. 3419-3424.
96. Pollmann, C., et al., *The Constitutive Photomorphogenesis 9 Signalosome Directs Vascular Endothelial Growth Factor Production in Tumor Cells*. Cancer Res, 2001. **61**: p. 8416-8421.
97. Kob, R., et al., *Regulation of the anaphase-promoting complex by the COP9 signalosome*. Cell Cycle, 2009. **8**: p. 2041-2049.
98. Leal, J.F., et al., *Cellular senescence bypass screen identifies new putative tumor suppressor genes*. Oncogene, 2008. **27**(14): p. 1961-70.

99. Yoneda-Kato, N., et al., *Myeloid leukemia factor 1 regulates p53 by suppressing COP1 via COP9 signalosome subunit 3*. The EMBO Journal, 2006. **24**: p. 1739–1749
100. Luo, J., et al., *A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene*. Cell, 2009. **137**(5): p. 835-48.
101. Boutonnet, C., et al., *Phosphorylation of Ser72 does not regulate the ubiquitin ligase activity and subcellular localization of Skp2*. Cell Cycle, 2010. **9**: p. 975-979.
102. Francis, J.M., et al., *Somatic mutation of CDKN1B in small intestine neuroendocrine tumors*. Nat Genet, 2013. **45**(12): p. 1483-6.
103. Stephens, P.J., et al., *The landscape of cancer genes and mutational processes in breast cancer*. Nature, 2012. **486**(7403): p. 400-4.
104. Carrano, C.A., et al., *SKP2 is required for ubiquitin-mediated degradation of the CDK inhibitor p27*. Nat Cell Biol, 1999. **1**: p. 193-199.
105. Esteva, J.F., et al., *Jun Activation Domain Binding Protein 1 Expression Is Associated with Low p27Kip1 Levels in Node-Negative Breast Cancer*. Clinical Cancer Research, 2003. **9**: p. 5652–5659.
106. Sui, L., et al., *Jab1 Expression Is Associated with Inverse Expression of p27kip1 and Poor Prognosis in Epithelial Ovarian Tumors*. Clinical Cancer Research, 2001. **7**: p. 4130–4135.
107. Tomoda, K., Y. Kubota, and J.Y. Kato, *Degradation of the cyclin-dependent-kinase inhibitor p27Kip1 is instigated by Jab1*. Nature, 1999. **398**: p. 160-165.
108. Yang, H.Y., et al., *Modified p27 Kip1 is efficient in suppressing HER2-mediated tumorigenicity*. J Cell Biochem, 2006. **98**(1): p. 128-38.
109. Kandoth, C., et al., *Mutational landscape and significance across 12 major cancer types*. Nature, 2013. **502**(7471): p. 333-9.
110. Tomoda, K., et al., *Multiple functions of Jab1 are required for early embryonic development and growth potential in mice*. J Biol Chem, 2004. **279**(41): p. 43013-8.
111. Zhou, B.P., et al., *HER-2/neu induces p53 ubiquitination via Akt-mediated MDM2 phosphorylation*. Nat Cell Biol, 2001. **3**: p. 973-892.
112. Bernardi, R., et al., *PML regulates p53 stability by sequestering Mdm2 to the nucleolus*. Nat Cell Biol, 2004. **6**(7): p. 665-72.
113. Oh, W., et al., *Jab1 induces the cytoplasmic localization and degradation of p53 in coordination with Hdm2*. J Biol Chem, 2006. **281**(25): p. 17457-65.
114. Tschumperlin, D.J., et al., *Mechanotransduction through growth-factor shedding into the extracellular space*. Nature, 2004. **429**(6987): p. 83-6.
115. Choi, H.H., et al., *COP9 signalosome subunit 6 stabilizes COP1, which functions as an E3 ubiquitin ligase for 14-3-3sigma*. Oncogene, 2011. **30**(48): p. 4791-801.
116. Chen, B., et al., *CDK inhibitor p57 (Kip2) is negatively regulated by COP9 signalosome subunit 6*. Cell Cycle, 2012. **11**(24): p. 4633-41.
117. Lee, M.H., I. Reynisdóttir, and J. Massagué, *Cloning of p57KIP2, a cyclin-dependent kinase inhibitor with unique domain structure and tissue distribution*. Genes Dev., 1995. **9**: p. 639-649.
118. Kamura, T., et al., *Degradation of p57Kip2 mediated by SCFSkp2-dependent ubiquitylation*. Proc Natl Acad Sci U S A, 2003. **100**(18): p. 10231-6.

119. Seeger, M., et al., *A novel protein complex involved in signal transduction possessing similarities to 26S proteasome subunits*. The FASEB journal, 1998. **12**: p. 469-478.
120. Wei, N., et al., *The COP9 complex is conserved between plants and mammals and is related to the 26S proteasome regulatory complex*. Curr Biol, 1998. **8**: p. 919-922.
121. Schwechheimer, C., G. Serino, and X.W. Deng, *Multiple Ubiquitin Ligase-Mediated Processes Require COP9 Signalosome and AXR1 Function*. The Plant Cell, 2002. **14**: p. 2553-2563.
122. Hetfeld, J.K.B., D. Bech - Otschir, and W. Dubiel, *Purification Method of the COP9 Signalosome from Human Erythrocytes*. 2005. **398**: p. 481-491.
123. Claret, F.-X., et al., *A new group of conserved coactivators that increase the specificity of AP-1 transcription factors*. Nature, 1996. **383**(6599): p. 453-457.
124. Bianchi, E., et al., *Integrin LFA-1 interacts with the transcriptional co-activator JAB1 to modulate AP-1 activity*. Nature, 2000. **404**: p. 617-621.
125. Kwok, F.S., et al., *Arabidopsis Homologs of a c-Jun Coactivator Are Present Both in Monomeric Form and in the COP9 Complex, and Their Abundance Is Differentially Affected by the Pleiotropic cop/det/fus Mutations*. The Plant Cell, 1998. **10**: p. 1779-1790.
126. Tomoda, K., et al., *The cytoplasmic shuttling and subsequent degradation of p27Kip1 mediated by Jab1/CSN5 and the COP9 signalosome complex*. J Biol Chem, 2002. **277**(3): p. 2302-10.
127. Fukumoto, A., et al., *Small Jab1-containing subcomplex is regulated in an anchorage- and cell cycle-dependent manner, which is abrogated by ras transformation*. FEBS Lett, 2005. **579**(5): p. 1047-54.
128. Sharon, M., et al., *Symmetrical modularity of the COP9 signalosome complex suggests its multifunctionality*. Structure, 2009. **17**(1): p. 31-40.
129. Wei, N. and X.W. Deng, *The COP9 signalosome*. Annu Rev Cell Dev Biol, 2003. **19**: p. 261-86.
130. Olma, M.H., et al., *An interaction network of the mammalian COP9 signalosome identifies Dda1 as a core subunit of multiple Cul4-based E3 ligases*. J Cell Sci, 2009. **122**(Pt 7): p. 1035-44.
131. Lykke-Andersen, K., et al., *Disruption of the COP9 Signalosome Csn2 Subunit in Mice Causes Deficient Cell Proliferation, Accumulation of p53 and Cyclin E, and Early Embryonic Death*. Molecular and Cellular Biology, 2003. **23**(19): p. 6790-6797.
132. Yan, J., et al., *COP9 Signalosome Subunit 3 Is Essential for Maintenance of Cell Proliferation in the Mouse Embryonic Epiblast*. Mol. Cell. Biol, 2003. **23**: p. 6798-6808.
133. Menon, S., et al., *COP9 signalosome subunit 8 is essential for peripheral T cell homeostasis and antigen receptor-induced entry into the cell cycle from quiescence*. Nat Immunol, 2007. **8**(11): p. 1236-45.
134. von Arnim, A.G. and C. Schwechheimer, *Life Is Degrading—Thanks to Some Zomes*. Molecular Cell, 2006. **23**(5): p. 621-629.
135. Wei, N., et al., *Arabidopsis COP8, COPIO, and COPII Genes Are Involved in Repression of Photomorphogenic Development in Darkness*. The Plant Cell, 1994. **6**: p. 629-643.

136. Gusmaroli, G., et al., *Role of the MPN subunits in COP9 signalosome assembly and activity, and their regulatory interaction with Arabidopsis Cullin3-based E3 ligases*. Plant Cell, 2007. **19**(2): p. 564-81.
137. Gusmaroli, G., S. Feng, and X.W. Deng, *The Arabidopsis CSN5A and CSN5B subunits are present in distinct COP9 signalosome complexes, and mutations in their JAMM domains exhibit differential dominant negative effects on development*. Plant Cell, 2004. **16**(11): p. 2984-3001.
138. Dohmann, E.M., C. Kuhnle, and C. Schwechheimer, *Loss of the CONSTITUTIVE PHOTOMORPHOGENIC9 signalosome subunit 5 is sufficient to cause the cop/det/fus mutant phenotype in Arabidopsis*. Plant Cell, 2005. **17**(7): p. 1967-78.
139. Jin, D., et al., *Plant COP9 Signalosome subunit 5, CSN5*. Plant Science, 2014. **224**: p. 54-61.
140. He, Q., et al., *The COP9 signalosome regulates the Neurospora circadian clock by controlling the stability of the SCFFWD-1 complex*. Genes Dev, 2005. **19**(13): p. 1518-31.
141. Zhou, Z., et al., *Neurospora COP9 signalosome integrity plays major roles for hyphal growth, conidial development, and circadian function*. PLoS Genet, 2012. **8**(5): p. e1002712.
142. Maytal-Kivity, V., et al., *The COP9 signalosome-like complex in S. cerevisiae and links to other PCI complexes*. The International Journal of Biochemistry & Cell Biology, 2003. **35**(5): p. 706-715.
143. Pick, E., et al., *The minimal deneddylase core of the COP9 signalosome excludes the Csn6 MPN- domain*. PLoS One, 2012. **7**(8): p. e43980.
144. Mundt, K.E., C. Liu, and A.M. Carr, *Deletion mutants in COP9/signalosome subunits in fission yeast Schizosaccharomyces pombe display distinct phenotypes*. Mol Biol Cell, 2002. **13**(2): p. 493-502.
145. V. Maytal-Kivity, et al., *COP9 signalosome components play a role in the mating pheromone response of S. cerevisiae*. EMBO Reports, 2002. **3**: p. 1215-1221.
146. Mundt, E.K., et al., *The COP9/signalosome complex is conserved in fission yeast and has a role in S phase*. Curr Biol, 1999. **9**: p. 1427-1430.
147. Busch, S., et al., *An eight-subunit COP9 signalosome with an intact JAMM motif is required for fungal fruit body formation*. Proc Natl Acad Sci U S A, 2007. **104**(19): p. 8089-94.
148. Komander, D., M.J. Clague, and S. Urbe, *Breaking the chains: structure and function of the deubiquitinases*. Nat Rev Mol Cell Biol, 2009. **10**(8): p. 550-63.
149. Nijman, S.M., et al., *A genomic and functional inventory of deubiquitinating enzymes*. Cell, 2005. **123**(5): p. 773-86.
150. Sowa, M.E., et al., *Defining the human deubiquitinating enzyme interaction landscape*. Cell, 2009. **138**(2): p. 389-403.
151. Wilkinson, K.D., *DUBs at a glance*. J Cell Sci, 2009. **122**(Pt 14): p. 2325-9.
152. Komander, D., et al., *Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains*. EMBO Rep, 2009. **10**(5): p. 466-73.
153. Bremm, A., S.M. Freund, and D. Komander, *Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne*. Nat Struct Mol Biol, 2010. **17**(8): p. 939-47.

154. Komander, D., *The emerging complexity of protein ubiquitination*. Biochem Soc Trans, 2009. **37**(Pt 5): p. 937-53.
155. Gong, L., *Identification of a Novel Isopeptidase with Dual Specificity for Ubiquitin- and NEDD8-conjugated Proteins*. Journal of Biological Chemistry, 2000. **275**(19): p. 14212-14216.
156. Catic, A., et al., *Screen for ISG15-crossreactive deubiquitinases*. PLoS One, 2007. **2**(7): p. e679.
157. Clague, M.J., J.M. Coulson, and S. Urbe, *Cellular functions of the DUBs*. J Cell Sci, 2012. **125**(Pt 2): p. 277-86.
158. Pathare, G.R., et al., *Crystal structure of the proteasomal deubiquitylation module Rpn8-Rpn11*. Proc Natl Acad Sci U S A, 2014. **111**(8): p. 2984-9.
159. Verma, R., et al., *Role of Rpn11 Metalloprotease in Deubiquitination and Degradation by the 26S Proteasome*. Science, 2002. **298**: p. 611-615.
160. Worden, E.J., C. Padovani, and A. Martin, *Structure of the Rpn11-Rpn8 dimer reveals mechanisms of substrate deubiquitination during proteasomal degradation*. Nat Struct Mol Biol, 2014. **21**(3): p. 220-7.
161. Yao, T. and R.E. Cohen, *A cryptic protease couples deubiquitination and degradation by the proteasome*. Nature, 2002. **419**(6905): p. 403-407.
162. Lam, Y., et al., *Editing of ubiquitin conjugates by an isopeptidase in the 26S proteasome*. 1997(Nature): p. 737-740.
163. Schwechheimer, C., *The COP9 signalosome (CSN): an evolutionary conserved proteolysis regulator in eukaryotic development*. Biochim Biophys Acta, 2004. **1695**(1-3): p. 45-54.
164. Doronkin, S., I. Djagaeva, and K.S. Beckendorf, *The COP9 Signalosome Promotes Degradation of Cyclin E during Early Drosophila Oogenesis*. Developmental Cell, 2003. **4**: p. 699-710.
165. Pintard, L., et al., *Neddylation and Deneddylation of CUL-3 Is Required to Target MEI-1/Katanin for Degradation at the Meiosis-to-Mitosis Transition in C. elegans*. Curr Biol, 2003. **13**: p. 911-921.
166. Chamovitz, A.D., et al., *The COP9 Complex, a Novel Multisubunit Nuclear Regulator Involved in Light Control of a Plant Developmental Switch*. Cell, 1996. **86**: p. 115-121.
167. Freilich, S., et al., *The COP9 signalosome is essential for development of Drosophila melanogaster*. Current Biology, 1999. **9**: p. 1187-1190.
168. Patterson-Fortin, J., et al., *Differential regulation of JAMM domain deubiquitinating enzyme activity within the RAP80 complex*. J Biol Chem, 2010. **285**(40): p. 30971-81.
169. Maytal-Kivity V., R.N., Hofmann K., Glickman M. H. , *MPN+, a putative catalytic motif found in a subset of MPN domain proteins from eukaryotes and prokaryotes, is critical for Rpn11 function*. BMC Biochemistry, 2002. **3:28**: p. 1471-2091/.
170. Rinaldi, T., et al., *A Mutation in a Novel Yeast Proteasomal Gene, RPN11/MPR1, Produces a Cell Cycle Arrest, Overreplication of Nuclear and Mitochondrial DNA, and an Altered Mitochondrial Morphology*. Molecular Biology of the Cell, 1998. **9**: p. 2917-2931.
171. Lee, M.J., et al., *Trimming of ubiquitin chains by proteasome-associated deubiquitinating enzymes*. Mol Cell Proteomics, 2011. **10**(5): p. R110 003871.



172. Bohn, S., et al., *Structure of the 26S proteasome from Schizosaccharomyces pombe at subnanometer resolution*. Proc Natl Acad Sci U S A, 2010. **107**(49): p. 20992-7.
173. da Fonseca, P.C., J. He, and E.P. Morris, *Molecular model of the human 26S proteasome*. Mol Cell, 2012. **46**(1): p. 54-66.
174. Lander, G.C., et al., *Complete subunit architecture of the proteasome regulatory particle*. Nature, 2012. **482**(7384): p. 186-91.
175. Lasker, K., et al., *Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach*. Proc Natl Acad Sci U S A, 2012. **109**(5): p. 1380-7.
176. Sanches, M., et al., *The crystal structure of the human Mov34 MPN domain reveals a metal-free dimer*. J Mol Biol, 2007. **370**(5): p. 846-55.
177. Jackson, R.J., C.U. Hellen, and T.V. Pestova, *The mechanism of eukaryotic translation initiation and principles of its regulation*. Nat Rev Mol Cell Biol, 2010. **11**(2): p. 113-27.
178. Zhou, M., et al., *Mass spectrometry reveals modularity and a complete subunit interaction map of the eukaryotic translation factor eIF3*. Proc Natl Acad Sci U S A, 2008. **105**(47): p. 18139-44.
179. Moretti, J., et al., *The translation initiation factor 3f (eIF3f) exhibits a deubiquitinase activity regulating Notch activation*. PLoS Biol, 2010. **8**(11): p. e1000545.
180. Cooper, E.M., J.D. Boeke, and R.E. Cohen, *Specificity of the BRISC deubiquitinating enzyme is not due to selective binding to Lys63-linked polyubiquitin*. J Biol Chem, 2010. **285**(14): p. 10344-52.
181. Shao, G., et al., *The Rap80-BRCC36 de-ubiquitinating enzyme complex antagonizes RNF8-Ubc13-dependent ubiquitination events at DNA double strand breaks*. Proc Natl Acad Sci U S A, 2009. **106**(9): p. 3166-71.
182. Cooper, E.M., et al., *K63-specific deubiquitination by two JAMM/MPN  $\beta$  complexes: BRISC-associated Brcc36 and proteasomal Poh1*. The EMBO Journal, 2009. **28**: p. 621-631.
183. Wang, B., et al., *NBA1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control*. Genes Dev, 2009. **23**(6): p. 729-39.
184. Shao, G., et al., *MERIT40 controls BRCA1-Rap80 complex integrity and recruitment to DNA double-strand breaks*. Genes Dev, 2009. **23**(6): p. 740-54.
185. Greenberg, R.A., *Recognition of DNA double strand breaks by the BRCA1 tumor suppressor network*. Chromosoma, 2008. **117**(4): p. 305-17.
186. Sato, Y., et al., *Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains*. Nature, 2008. **455**(7211): p. 358-62.
187. Kikuchi, K., et al., *Identification of AMSH-LP containing a Jab1/MPN domain metalloenzyme motif*. Biochemical and Biophysical Research Communications, 2003. **306**: p. 637-643.
188. Shrestha, R.K., et al., *Insights into the mechanism of deubiquitination by JAMM deubiquitinases from co-crystal structures of enzyme with substrate and product*. Biochemistry, 2014.
189. Shen, L.N., et al., *Structural basis of NEDD8 ubiquitin discrimination by the deNEDDylating enzyme NEDP1*. EMBO J, 2005. **24**(7): p. 1341-51.



190. Crow, A., et al., *The molecular basis of ubiquitin-like protein NEDD8 deamidation by the bacterial effector protein Cif*. Proc Natl Acad Sci U S A, 2012. **109**(27): p. E1830-8.
191. Li, T., et al., *High-level expression and purification of recombinant SCF ubiquitin ligases*. Methods Enzymol, 2005. **398**: p. 125-42.
192. Winn, M.D., et al., *Overview of the CCP4 suite and current developments*. Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 235-42.
193. Langer, G., et al., *Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7*. Nat Protoc, 2008. **3**(7): p. 1171-9.
194. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis*. J Comput Chem, 2004. **25**(13): p. 1605-12.
195. Eswar, N., et al., *Comparative protein structure modeling using MODELLER*. Curr Protoc Protein Sci, 2007. **Chapter 2**: p. Unit 2 9.
196. Pang, Y.P., et al., *Successful molecular dynamics simulation of the zinc-bound farnesyltransferase using the cationic dummy atom approach*. Protein Sci, 2000. **9**(10): p. 1857-65.
197. Case, D.A., *AMBER 11*. 2011.
198. Duan, Y., et al., *A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations*. J Comput Chem, 2003. **24**(16): p. 1999-2012.
199. Bart, E., Kuczera, K., Leimkuhler, B., Skeel, R., *Algorithms for constrained molecular dynamics*. J. Comput. Chem, 1995. **16**: p. 1192-1209.
200. Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., Pedersen, L.G., *A smooth particle mesh Ewald method*. J. Chem. Phys, 1995. **103**: p. 8577-8593.
201. Case, D.A., et al., *The Amber biomolecular simulation programs*. J Comput Chem, 2005. **26**(16): p. 1668-88.
202. Piotto, M., V. Saudek, and V. Sklenar, *Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions*. J Biomol NMR, 1992. **2**(6): p. 661-5.
203. Sklenar, V., et al., *Two- and three-dimensional HCN experiments for correlating base and sugar resonances in <sup>15</sup>N,<sup>13</sup>C-labeled RNA oligonucleotides*. J Biomol NMR, 1993. **3**(6): p. 721-7.
204. Malliavin, T.E., J.L. Pons, and M.A. Delsuc, *An NMR assignment module implemented in the Gifa NMR processing program*. Bioinformatics, 1998. **14**(7): p. 624-31.
205. Englander, S.W. and A.J. Wand, *Main-chain-directed strategy for the assignment of <sup>1</sup>H NMR spectra of proteins*. Biochemistry, 1987. **26**(19): p. 5953-8.
206. Wand, A.J. and S.J. Nelson, *Refinement of the main chain directed assignment strategy for the analysis of <sup>1</sup>H NMR spectra of proteins*. Biophys J, 1991. **59**(5): p. 1101-12.
207. Roth, S.M., et al., *Characterization of the secondary structure of calmodulin in complex with a calmodulin-binding domain peptide*. Biochemistry, 1992. **31**(5): p. 1443-51.
208. Fiser, A. and A. Sali, *Modeller: generation and refinement of homology-based protein structure models*. Methods Enzymol, 2003. **374**: p. 461-91.

209. Lawrence, M.C. and P.M. Colman, *Shape complementarity at protein/protein interfaces*. J Mol Biol, 1993. **234**(4): p. 946-50.
210. Krissinel, E. and K. Henrick, *Inference of macromolecular assemblies from crystalline state*. J Mol Biol, 2007. **372**(3): p. 774-97.
211. Gracy, J. and L. Chiche, *PAT: a protein analysis toolkit for integrated biocomputing on the web*. Nucleic Acids Res, 2005. **33**(Web Server issue): p. W65-71.
212. Dubiel, T.S.a.W., *Conjugation and Deconjugation of Ubiquitin Family Modifiers*. CONTROL OF DENEDDYLATION BY THE COP9 SIGNALOSOME, ed. M. Groettrup. 2010.
213. Wolberger, C., *Mechanisms for regulating deubiquitinating enzymes*. Protein Sci, 2014. **23**(4): p. 344-53.
214. Mayer, A.N. and K.D. Wilkinson, *Detection, Resolution, and Nomenclature of Multiple Ubiquitin Carboxyl-Terminal Esterases from Bovine Calf Thymus*. Biochemistry, 1988. **28**: p. 166-172.
215. Hassink, G.C., et al., *The ER-resident ubiquitin-specific protease 19 participates in the UPR and rescues ERAD substrates*. EMBO Rep, 2009. **10**(7): p. 755-61.
216. Nakamura, N. and S. Hirose, *Regulation of mitochondrial morphology by USP30, a deubiquitinating enzyme present in the mitochondrial outer membrane*. Mol Biol Cell, 2008. **19**(5): p. 1903-11.
217. Zhu, X., R. Menard, and T. Sulea, *High incidence of ubiquitin-like domains in human ubiquitin-specific proteases*. Proteins, 2007. **69**(1): p. 1-7.
218. Sun, S.C., *CYLD: a tumor suppressor deubiquitinase regulating NF-kappaB activation and diverse biological processes*. Cell Death Differ, 2010. **17**(1): p. 25-34.
219. Reyes-Turcu, F.E., K.H. Ventii, and K.D. Wilkinson, *Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes*. Annu Rev Biochem, 2009. **78**: p. 363-97.
220. Min Hu, P.L., Muyang Li, Wenyu Li, Tingting Yao, Jia-Wei Wu, Wei Gu, and a.Y.S. Robert E. Cohen, *Crystal Structure of a UBP-Family Deubiquitinating Enzyme in Isolation and in Complex with Ubiquitin Aldehyde*. 2002. **111**: p. 1041-1054.
221. Reyes-Turcu, F.E., et al., *Recognition of polyubiquitin isoforms by the multiple ubiquitin binding modules of isopeptidase T*. J Biol Chem, 2008. **283**(28): p. 19581-92.
222. Min Hu, P.L., Ling Song, Philip D Jeffrey, Tatiana A Chernova, Keith D Wilkinson, Robert E Cohen and Yigong Shi, *Structure and mechanisms of the proteasome- associated deubiquitinating enzyme USP14*. The EMBO Journal, 2005. **24**: p. 3747-3756
223. Komander, D., et al., *The structure of the CYLD USP domain explains its specificity for Lys63-linked polyubiquitin and reveals a B box module*. Mol Cell, 2008. **29**(4): p. 451-64.
224. Fang, Y., D. Fu, and X.Z. Shen, *The potential role of ubiquitin c-terminal hydrolases in oncogenesis*. Biochim Biophys Acta, 2010. **1806**(1): p. 1-6.
225. Popp, M.W., K. Artavanis-Tsakonas, and H.L. Ploegh, *Substrate filtering by the active site crossover loop in UCHL3 revealed by sortagging and gain-of-function mutations*. J Biol Chem, 2009. **284**(6): p. 3593-602.

226. Larsen, C.N., B.A. Krantz, and K.D. Wilkinson, *Substrate Specificity of Deubiquitinating Enzymes: Ubiquitin C-Terminal Hydrolases*. *Biochemistry*, 1998. **37**: p. 3358-3368.
227. Steven C. Johnston, S.M.R., Robert E. Cohen and Christopher P. Hill, *Structural basis for the specificity of ubiquitin C-terminal hydrolases*. *The EMBO Journal* 1999. **18**: p. 3877–3887.
228. Finley, D., *Recognition and processing of ubiquitin-protein conjugates by the proteasome*. *Annu Rev Biochem*, 2009. **78**: p. 477-513.
229. Nanao, M.H., et al., *Crystal structure of human otubain 2*. *EMBO Rep*, 2004. **5**(8): p. 783-8.
230. Komander, D. and D. Barford, *Structure of the A20 OTU domain and mechanistic insights into deubiquitination*. *Biochem J*, 2008. **409**(1): p. 77-85.
231. Edelmann, M.J., et al., *Structural basis and specificity of human otubain 1-mediated deubiquitination*. *Biochem J*, 2009. **418**(2): p. 379-90.
232. Denu, J.M. and K.G. Tanner, *Specific and Reversible Inactivation of Protein Tyrosine Phosphatases by Hydrogen Peroxide: Evidence for a Sulfenic Acid Intermediate and Implications for Redox Regulation*. *Biochemistry*, 1998. **37**: p. 5633-5642.
233. Enesa, K., et al., *Hydrogen peroxide prolongs nuclear localization of NF-kappaB in activated cells by suppressing negative regulatory mechanisms*. *J Biol Chem*, 2008. **283**(27): p. 18582-90.
234. Kulathu, Y., et al., *Regulation of A20 and other OTU deubiquitinases by reversible oxidation*. *Nat Commun*, 2013. **4**: p. 1569.
235. Mao, Y., et al., *Deubiquitinating function of ataxin-3: insights from the solution structure of the Josephin domain*. *Proc Natl Acad Sci U S A*, 2005. **102**(36): p. 12700-5.
236. Nicastro, G., et al., *The solution structure of the Josephin domain of ataxin-3: structural determinants for molecular recognition*. *Proc Natl Acad Sci U S A*, 2005. **102**(30): p. 10493-8.
237. Huong J. T. T. Tran, M.D.A., Jan Löwe, Mark Bycroft, *Structure of the Jab1/MPN Domain and Its Implications for Proteasome Function*. *Biochemistry*, 2003. **42**: p. 11460-11465.
238. Zhu, P., et al., *A histone H2A deubiquitinase complex coordinating histone acetylation and H1 dissociation in transcriptional regulation*. *Mol Cell*, 2007. **27**(4): p. 609-21.
239. Reiley, W., et al., *Regulation of the deubiquitinating enzyme CYLD by IkkappaB kinase gamma-dependent phosphorylation*. *Mol Cell Biol*, 2005. **25**(10): p. 3886-95.
240. Meulmeester, E., et al., *Mechanism and consequences for paralog-specific sumoylation of ubiquitin-specific protease 25*. *Mol Cell*, 2008. **30**(5): p. 610-9.
241. Todi, S., V; Winborn, B, J; Scaglione, K, M; Blount, J, R; Travis, S, M; Paulson, H, L, *Ubiquitination directly enhances activity of the deubiquitinating enzyme ataxin-3*. *The EMBO Journal*, 2009. **28**(4): p. 372–382.
242. Parry, G. and M. Estelle, *Regulation of cullin-based ubiquitin ligases by the Nedd8/RUB ubiquitin-like proteins*. *Seminars in Cell & Developmental Biology*, 2004. **15**(2): p. 221-229.

243. Duda, D.M., et al., *Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation*. Cell, 2008. **134**(6): p. 995-1006.
244. McCullough, J., et al., *Activation of the endosome-associated ubiquitin isopeptidase AMSH by STAM, a component of the multivesicular body-sorting machinery*. Curr Biol, 2006. **16**(2): p. 160-5.
245. Row, P.E., et al., *The ubiquitin isopeptidase UBPY regulates endosomal ubiquitin dynamics and is essential for receptor down-regulation*. J Biol Chem, 2006. **281**(18): p. 12618-24.
246. Berlin, I., H. Schwartz, and P.D. Nash, *Regulation of epidermal growth factor receptor ubiquitination and trafficking by the USP8.STAM complex*. J Biol Chem, 2010. **285**(45): p. 34909-21.
247. Endo, A., et al., *Nucleolar structure and function are regulated by the deubiquitylating enzyme USP36*. J Cell Sci, 2009. **122**(Pt 5): p. 678-86.
248. Yao, T., et al., *Distinct modes of regulation of the Uch37 deubiquitinating enzyme in the proteasome and in the Ino80 chromatin-remodeling complex*. Mol Cell, 2008. **31**(6): p. 909-17.
249. Reyes-Turcu, F.E., et al., *The ubiquitin binding domain ZnF UBP recognizes the C-terminal diglycine motif of unanchored ubiquitin*. Cell, 2006. **124**(6): p. 1197-208.
250. Kim, M.S., et al., *STAM-AMSH interaction facilitates the deubiquitination activity in the C-terminal AMSH*. Biochem Biophys Res Commun, 2006. **351**(3): p. 612-8.
251. Kato, J.Y. and N. Yoneda-Kato, *Mammalian COP9 signalosome*. Genes Cells, 2009. **14**(11): p. 1209-25.
252. Huang, J., et al., *Jab1 mediates protein degradation of the Rad9-Rad1-Hus1 checkpoint complex*. J Mol Biol, 2007. **371**(2): p. 514-27.
253. Dawadschargal Bech - Otschir, et al., *COP9 signalosome - specific phosphorylation targets p53 to degradation by the ubiquitin system*. The EMBO Journal, 2001. **20**(7): p. 1630-1639.
254. S F Kwok, et al., *Arabidopsis homologs of a c-Jun coactivator are present both in monomeric form and in the COP9 complex, and their abundance is differentially affected by the pleiotropic cop/det/fus mutations*. The Pant Cell, 1998. **10**: p. 1779-1790.
255. Maarten Fornerod, et al., *CRM1 Is an Export Receptor for Leucine-Rich Nuclear Export Signals*. Cell, 1997. **90**: p. 1051-1060.
256. Chamovitz , D.A.S., D, *JAB1/CSN5 and the COP9 signalosome*. EMBO reports, 2001. **21**: p. 96-101
257. Nguyen, T.H., et al., *Structural basis of Brr2-Prp8 interactions and implications for U5 snRNP biogenesis and the spliceosome active site*. Structure, 2013. **21**(6): p. 910-19.
258. Beck, F., et al., *Near-atomic resolution structural model of the yeast 26S proteasome*. Proc Natl Acad Sci U S A, 2012. **109**(37): p. 14870-5.
259. Kotiguda, G.G., et al., *The organization of a CSN5-containing subcomplex of the COP9 signalosome*. J Biol Chem, 2012. **287**(50): p. 42031-41.
260. Kelley, L.A. and M.J. Sternberg, *Protein structure prediction on the Web: a case study using the Phyre server*. Nat Protoc, 2009. **4**(3): p. 363-71.

261. Davies, C.W., et al., *Structural and thermodynamic comparison of the catalytic domain of AMSH and AMSH-LP: nearly identical fold but different stability*. J Mol Biol, 2011. **413**(2): p. 416-29.
262. O'Donoghue, J.E., et al., *Nedd8 processing enzymes in Schizosaccharomyces pombe*. BMC Biochem, 2013. **14**: p. 8.
263. Mullally, J.E., et al., *Cyclopentenone prostaglandins of the J series inhibit the ubiquitin isopeptidase activity of the proteasome pathway*. J Biol Chem, 2001. **276**(32): p. 30366-73.
264. Hurley, J.H., S. Lee, and G. Prag, *Ubiquitin-binding domains*. Biochem J, 2006. **399**(3): p. 361-72.
265. Bellare, P., et al., *Ubiquitin binding by a variant Jab1/MPN domain in the essential pre-mRNA splicing factor Prp8p*. RNA, 2006. **12**(2): p. 292-302.
266. Rockel, B., et al., *Electron microscopy and in vitro deneddylation reveal similar architectures and biochemistry of isolated human and Flag-mouse COP9 signalosome complexes*. Biochem Biophys Res Commun, 2014.
267. Querol-Audi, J., et al., *Architecture of human translation initiation factor 3*. Structure, 2013. **21**(6): p. 920-8.
268. Yoshida, A., N. Yoneda-Kato, and J.Y. Kato, *CSN5 specifically interacts with CDK2 and controls senescence in a cytoplasmic cyclin E-mediated manner*. Sci Rep, 2013. **3**: p. 1054.
269. Fuzesi-Levi, M.G., et al., *Dynamic regulation of the COP9 signalosome in response to DNA damage*. Mol Cell Biol, 2014. **34**(6): p. 1066-76.
270. Gastaldello, S., et al., *A deneddylase encoded by Epstein-Barr virus promotes viral DNA replication by regulating the activity of cullin-RING ligases*. Nat Cell Biol, 2010. **12**(4): p. 351-61.
271. Hicke, L., H.L. Schubert, and C.P. Hill, *Ubiquitin-binding domains*. Nat Rev Mol Cell Biol, 2005. **6**(8): p. 610-21.
272. Haas, A.L. and P. Bright, *The Dynamics of Ubiquitin Pools within Cultured Human Lung Fibroblasts*. The Journal of Biological Chemistry, 1987. **262**(1): p. 345-351.
273. Cuff, J.A. and G.J. Barton, *Application of Multiple Sequence Alignment Profiles to Improve Protein Secondary Structure Prediction*. PROTEINS: Structure, Function and Genetics 2000. **40**: p. 502-511.
274. Zhang, L., et al., *Crystal structure of the C-terminal domain of splicing factor Prp8 carrying retinitis pigmentosa mutants*. Protein Sci, 2007. **16**(6): p. 1024-31.
275. Solomons, J., et al., *Structural basis for ESCRT-III CHMP3 recruitment of AMSH*. Structure, 2011. **19**(8): p. 1149-59.
276. Birol, M., et al., *Structural and Biochemical Characterization of the Cop9 Signalosome CSN5/CSN6 Heterodimer*. PLoS One, 2014. **9**(8): p. e105688.
277. Enchev, R.I., et al., *Structural insights into the COP9 signalosome and its common architecture with the 26S proteasome lid and eIF3*. Structure, 2010. **18**(4): p. 518-27.
278. Zuiderweg, E.R.P., *Mapping Protein-Protein Interactions in Solution by NMR Spectroscopy*. Biochemistry, 2002. **41**: p. 1-7.
279. Grishin, N.V. and M.A. Phillips, *The subunit interfaces of oligomeric enzymes are conserved to a similar extent to the overall protein sequences*. Protein Science, 1994. **3**: p. 2455-2458.



280. Chung, J.L., W. Wang, and P.E. Bourne, *Exploiting sequence and structure homologs to identify protein-protein binding sites*. *Proteins*, 2006. **62**(3): p. 630-40.
281. Lin-nan Shen, et al., *Structural basis of NEDD8 ubiquitin discrimination by the deNEDDylating enzyme NEDP1*. *The EMBO Journal*, 2005. **24**: p. 1341-1351.
282. Kamitani, T., et al., *Characterization of NEDD8, a Developmentally Down-regulated Ubiquitin-like Protein*. *Journal of Biological Chemistry*, 1997. **272**(45): p. 28557-28562.
283. Osaka, F., et al., *Covalent modifier NEDD8 is essential for SCF ubiquitin-ligase in fission yeast*. *The EMBO Journal*, 2000. **19**: p. 3475-3484.
284. Callis, J., et al., *Structure and Evolution of Genes Encoding Polyubiquitin and Ubiquitin-Like Proteins in Arabidopsis thaliana Ecotype Columbia*. *Genetics*, 1995. **159**: p. 921-939.
285. Rao-Naik, C., et al., *The Rub Family of Ubiquitin-like Proteins Crystal structure of Arabidopsis Rub1 and expression of multiple rubs in Arabidopsis*. *J. Biol. Chem.*, 1998. **273**: p. 34976-34982.
286. Yu, H.A., et al., *Characterization of ubiquitin C-terminal hydrolase 1 (YUH1) from Saccharomyces cerevisiae expressed in recombinant Escherichia coli*. *Protein Expr Purif*, 2007. **56**(1): p. 20-6.
287. Christmann, M., et al., *Control of multicellular development by the physically interacting deneddylases DEN1/DenA and COP9 signalosome*. *PLoS Genet*, 2013. **9**(2): p. e1003275.
288. Chan, Y., et al., *DEN1 deneddylates non-cullin proteins in vivo*. *J Cell Sci*, 2008. **121**(Pt 19): p. 3218-23.
289. Kurihara, L.J., et al., *Expression and Functional Analysis of Uch-L3 during Mouse Development*. *Mol Biol Cell*, 2000. **20**: p. 2498-2504.
290. Ebina, M., et al., *Myeloma overexpressed 2 (Myeov2) regulates L11 subnuclear localization through Nedd8 modification*. *PLoS One*, 2013. **8**(6): p. e65285.
291. Watson, I.R., et al., *Chemotherapy induces NEDP1-mediated destabilization of MDM2*. *Oncogene*, 2010. **29**(2): p. 297-304.
292. Hallstrom, T.C. and J.R. Nevins, *Jab1 is a specificity factor for E2F1-induced apoptosis*. *Genes Dev*, 2006. **20**(5): p. 613-23.
293. Arnim, G.A. and X.W. Deng, *Light inactivation of Arabidopsis Photomorphogenic Repressor COP1 Involves a Cell-Specific Regulation of Its Nucleocytoplasmic Partitioning*. *Cell*, 1994. **79**: p. 1035-1045



# Insights into the regulation of the human COP9 signalosome catalytic subunit, CSN5/Jab1

Aude Echalié<sup>a,1</sup>, Yunbao Pan<sup>b,2</sup>, Melissa Birol<sup>a,2</sup>, Nicolas Tavernier<sup>c</sup>, Lionel Pintard<sup>c</sup>, François Hoh<sup>a</sup>, Christine Ebel<sup>d</sup>, Nathalie Galoppe<sup>a</sup>, François X. Claret<sup>b,e</sup>, and Christian Dumas<sup>a</sup>

<sup>a</sup>Centre de Biochimie Structurale, Centre National de la Recherche Scientifique (CNRS), Institut National de la Santé et de la Recherche Médicale-Unités Mixtes de Recherche 5048 and 1054, Université Montpellier I, 34090 Montpellier, France; <sup>b</sup>Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030; <sup>c</sup>Institut Jacques Monod, 75013 Paris, France; <sup>d</sup>Institut de Biologie Structurale, CNRS, Commissariat à l'Energie Atomique, Université Joseph Fourier, 38027 Grenoble, France; and <sup>e</sup>Cancer Biology Program and Experimental Therapeutic Program, The University of Texas Graduate School of Biomedical Sciences, Houston, TX 77030

Edited by Wolfgang Baumeister, Max Planck Institute of Biochemistry, Martinsried, Germany, and approved December 5, 2012  
(received for review June 4, 2012)

The COP9 (Constitutive photomorphogenesis 9) signalosome (CSN), a large multiprotein complex that resembles the 19S lid of the 26S proteasome, plays a central role in the regulation of the E3-cullin RING ubiquitin ligases (CRLs). The catalytic activity of the CSN complex, carried by subunit 5 (CSN5/Jab1), resides in the deneddylation of the CRLs that is the hydrolysis of the cullin-neural precursor cell expressed developmentally downregulated gene 8 (Nedd8) isopeptide bond. Whereas CSN-dependent CSN5 displays isopeptidase activity, it is intrinsically inactive in other physiologically relevant forms. Here we analyze the crystal structure of CSN5 in its catalytically inactive form to illuminate the molecular basis for its activation state. We show that CSN5 presents a catalytic domain that brings essential elements to understand its activity control. Although the CSN5 active site is catalytically competent and compatible with di-isopeptide binding, the Ins-1 segment obstructs access to its substrate-binding site, and structural rearrangements are necessary for the Nedd8-binding pocket formation. Detailed study of CSN5 by molecular dynamics unveils signs of flexibility and plasticity of the Ins-1 segment. These analyses led to the identification of a molecular trigger implicated in the active/inactive switch that is sufficient to impose on CSN5 an active isopeptidase state. We show that a single mutation in the Ins-1 segment restores biologically relevant deneddylase activity. This study presents detailed insights into CSN5 regulation. Additionally, a dynamic monomer-dimer equilibrium exists both in vitro and in vivo and may be functionally relevant.

cullin regulation | protein degradation | MPN | Rpn11

Cell signaling processes mediated by ubiquitylation, the posttranslational covalent conjugation of ubiquitin molecules, are of prime importance for cellular activity and particularly for protein turnover. Ubiquitin-ligase enzymes (E3s) are responsible for the last step of the ubiquitylation reaction, and the multi-subunit cullin-RING E3 ubiquitin ligases (CRLs) represent the most prominent of E3 enzymes. Among the several factors that regulate CRL activity, cullin neddylation/deneddylation cycles are central (1).

The Cop9 signalosome (CSN), which is an eight-subunit complex largely conserved through evolution, deneddylates CRLs and thereby regulates CRL activity. As a large number of proteins are ubiquitylated by CRLs, the CSN complex is implicated in the control of a significant proportion of the proteome, including protooncogenes, tumor suppressors, and other important cellular protagonists (1). Not surprisingly, the CSN has been implicated in various cellular functions, ranging from cell cycles to circadian rhythm and to immunity in various organisms. Furthermore, many studies have found a strong link between the CSN and cancers (2).

The CSN, a multi-protein complex of about 320 kDa, contains six proteasome Cop9 eIF3 (PCI)-based subunits and two MPN1-Pad1-N-terminal (MPN)-based subunits. The subunit 5 [CSN5; also known as c-Jun activation domain-binding protein-1 (Jab1)] (3), one of the two MPN-containing subunits, carries a zinc-

dependent isopeptidase catalytic center that contains a Jab1/MPN/Mov34 (JAMM) motif (also known as MPN<sup>+</sup> motif) (4). Several detailed studies suggested that the organization of the CSN complex resembles that of the 26S proteasome lid (5), with the deubiquitinase enzyme Rpn11 being the equivalent of the deneddylating subunit CSN5 (4, 6). The physiology of the CSN has been well researched (2). Intriguingly, the CSN cancer implication is attributable mainly to CSN5, which is located on human chromosome 8q.

Smaller forms of the holo-CSN complex, with variable compositions, have been found in vivo (7–11). Although possibly important in cell cycle progression, these sub-CSN complexes have not yet been fully functionally characterized (12). It is interesting that, as alluded to for Rpn11 in several reports (6, 13), CSN5 is found in two forms: a holo-CSN-associated form that is catalytically active and a stand-alone state void of isopeptidase activity (4, 5). The modularity and topology of the CSN complex have been explored in vitro by non-denaturing mass spectrometry (MS), which revealed that CSN5 is a peripheral subunit that can homo-dimerize outside of the CSN complex and interacts mostly with the other MPN-containing subunit, CSN6, in the context of the CSN complex (5). The potential interactions of CSN5 with other CSN subunits, namely CSN1, CSN2, CSN4, and CSN7, have also been highlighted (1, 8, 14, 15).

To elucidate the molecular regulation of CSN5 activity, we structurally and functionally characterized it in its CSN-independent form by X-ray crystallography, molecular dynamics (MD) simulations, and in vitro and in vivo studies. Our structural work uncovered a potential molecular trigger that regulates the active/inactive transition of CSN5. These experiments contributed to the design of a constitutively active form of isolated CSN5, shedding lights on its activation control mechanism at a molecular level.

## Results

**Overall Structure and Oligomeric Arrangement.** A stable form of human CSN5 comprising residues 1–257 (CSN5<sub>1–257</sub>), identified by MS and N-terminal sequencing, was isolated and crystallized. The crystal structure was solved by selenium single-wavelength anomalous dispersion (SAD) using diffraction data to 2.6 Å (SI

Author contributions: A.E., F.X.C., and C.D. designed research; A.E., Y.P., M.B., N.T., F.H., C.E., N.G., F.X.C., and C.D. performed research; A.E., Y.P., L.P., F.X.C., and C.D. contributed new reagents/analytic tools; A.E., L.P., F.X.C., and C.D. analyzed data; and A.E. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors reported in this paper have been deposited in the Protein Data Bank, [www.pdb.org](http://www.pdb.org) (PDB ID code 4F7O).

<sup>1</sup>To whom correspondence should be addressed. E-mail: aude.echalié-glazer@cbs.cnrs.fr.

<sup>2</sup>Y.P. and M.B. contributed equally to this work.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1209345110/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1209345110/-DCSupplemental).

Appendix, Table S1). CSN5, which consists of 334 residues, is a Jab1/MPN superfamily member with a conserved core MPN domain (51–230) and a JAMM motif (Glu76, His138, His140, and Asp151) (Fig. 1 *A* and *B*). In addition to the MPN catalytic domain, CSN5 possesses N- and C-terminal regions that tightly pack against the MPN fold and form an extended catalytic domain.

The asymmetric unit of CSN5<sub>1–257</sub> crystal contains a dimer, related by a local twofold axis perpendicular to the crystallographic dyad axis, also generating a second dimeric arrangement (SI Appendix; SI Appendix, Fig. S14). The evidence that CSN5 could adopt a homo-oligomeric arrangement in vivo could be of substantial interest and was further investigated.

**CSN5 Can Form Homo-Dimers Both In Vitro and In Vivo.** To investigate the presence of the oligomeric species, our experimental approach was based on chemical cross-linking (SI Appendix, Fig. S1B), dynamic light scattering (DLS; SI Appendix, Table S3), and analytical ultracentrifugation (AUC; SI Appendix, Table S4 and Fig. S1 *C* and *D*). The results showed that monomers and dimers were the major species of CSN5 detected in solution. Investigation of the presence of CSN5 oligomers in vivo carried out by coimmunoprecipitation experiments on mammalian cell extracts confirmed that CSN5 assembles in dimer (Fig. 1*C*). Supported by both in vitro and in vivo data, these observations

suggest that a CSN5 dimeric assembly could be present in solution in equilibrium with monomeric species (SI Appendix, Tables S3 and S4). It is noteworthy that other MPN-containing proteins were found to assemble in dimers in the crystals and that each of the described dimers proceeds via totally different interfaces (16, 17), preventing further comparison. Moreover, the question of the physiological relevance of these dimers has not yet been addressed in vivo.

In addition to these experiments and on the basis of the A–B and A–A' dimer interface analysis, mutations or deletions were designed to selectively weaken these two intersubunit interactions. These data confirmed that CSN5 dimers proceed both in vitro and in vivo predominantly via the A–B interface (SI Appendix; SI Appendix, Fig. S1, *E–G*) where more than 50% of the contributing residues are highly conserved among the 170 available sequences (SI Appendix, Fig. S2), further suggesting that this assembly may be physiologically relevant. Taken together, these results suggest that CSN5 could form dimers, unveiling a potential new level of regulation in the biology of CSN5. Additional studies will be needed to understand the biological role and the distribution of the CSN5 dimeric form in cells.

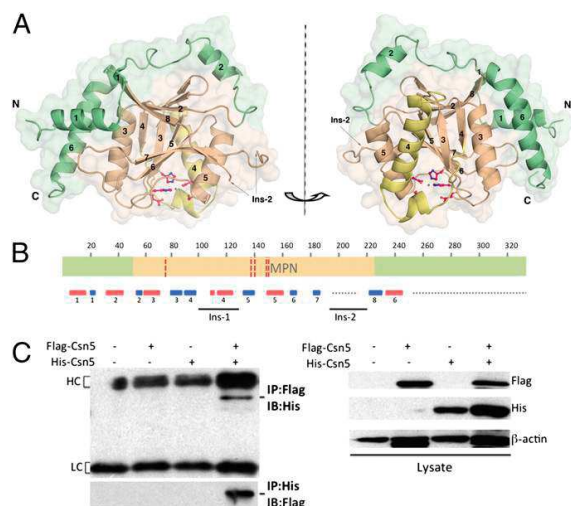
**Conserved Rigid MPN Domain Is Decorated by CSN5-Specific N- and C-Terminal Extensions.** The CSN5<sub>1–257</sub> structure reveals a fold typical to the Jab1/MPN superfamily that has been described in a number of structures (Fig. 1*A*; SI Appendix; SI Appendix, Fig. S3) (16–21). Interestingly structural comparison between MPN members revealed that the region spanning from residues 97–131 (referred to as Ins-1) displays an array of conformations in the various MPN members (Fig. 2*A*). A sequence alignment focused on the Ins-1 segment of four representatives of the MPN+/JAMM family is presented in Fig. 2*B*. It is noteworthy that the lack of electron density for the CSN5 portion consisting of residues 197–219 [corresponding to Ins-2 in the structure of associated molecule with the SH3 domain of STAM (signal transducing adapter molecule) deubiquitinase (AMSH-LP) (20)] prevented accurate modeling and analysis of this segment.

The ensemble of the CSN5-specific N- and C-terminal segments wraps around and makes extensive contacts with the conserved MPN domain core (Fig. 1*A*; SI Appendix, Fig. S3). Most MPN protein structures solved to date display reduced or are lacking N- and C-terminal additions, with the exception of Prp8p structure that has extensions of similar size to that of CSN5 (19). However, in CSN5 and in the Prp8p scaffolding protein, these regions adopt very different positions and conformations with respect to the core MPN domain.

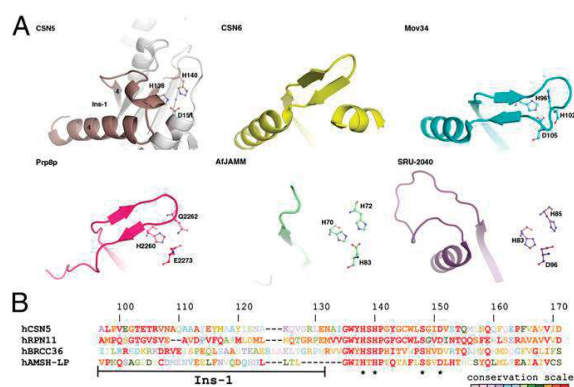
To extend and complement the structural insights obtained from crystallography, we carried out a series of MD simulations. The CSN5 crystal structure suggests that the central core domain is stable and that some anking  $\alpha$ -helices and loops displaying higher *B*-factors could be locked into the structure due to the crystal packing. MD simulations of the solvated CSN5 monomer at 300 K for 40 ns confirmed that the core domain is stable (SI Appendix, Fig. S4*A*) and that the residues forming the Ins-2 segment, the loops, and the N- and C-terminal ends display the maximum fluctuation compared with the central core domain (SI Appendix, Fig. S4*B*).

**CSN5 Zinc-Binding Site Is Catalytically Competent, Similar to Other JAMM-Containing Motifs.** As we anticipated from other MPN+/JAMM proteases, the CSN5 structure contains one zinc atom (Fig. 3*A*). The strictly conserved zinc coordination site (Fig. 2*B*) is composed of residues from helix  $\alpha 5$  and a subset of the central  $\beta$ -sheet ( $\beta 5$ ,  $\beta 5$ - $\alpha 5$ ,  $\beta 6$ , and  $\beta 7$ ). The zinc is tetrahedrally coordinated to two His residues (His138 and His140), one Asp residue (Asp151), and a catalytic water molecule hydrogen bonded to Glu76 and Ser148. The importance of the active site zinc-coordinating residues in catalysis was previously tested by mutagenesis (4).

To date, AMSH-LP is the only structural example of an active MPN+/JAMM isopeptidase enzyme that exists in an



**Fig. 1.** CSN5 overall structure and oligomeric arrangement. (*A*) Overall monomeric structure of CSN5<sub>1–257</sub> reveals a central MPN core domain (light brown) and peripheral N- and C-terminal extensions (green). The Ins-1 region (residues 97–131) is yellow. Secondary structure elements are numbered, and the catalytic center is shown in ball-and-stick representation. (*B*) Schematic representation of the CSN5 domain organization. The MPN domain is delineated in orange; the N- and C-terminal extensions are green. The five residues of the JAMM motif, namely E76, H138, H140, S148, and D151, are indicated by red dotted lines. The Ins-1 and Ins-2 insertions are placed with respect to their positions in the sequence. The secondary structure elements are shown in red for  $\alpha$ -helices and in blue for  $\beta$ -strands. Regions that are either not ordered in the crystal structure or not included in the crystallized fragment are indicated by a gray dotted line. (*C*) Biochemical identification of CSN5 dimer in vivo by coimmunoprecipitation experiments. 293T cells were transfected with either Flag- or His-tagged CSN5 or both for 48 h, and then cell lysates were immunoprecipitated (IP) with either anti-Flag (Upper) or anti-His (Lower) tag antibodies and immunoblotted (IB) with anti-His or anti-Flag tag antibodies, respectively (Right). The expression of the corresponding pairs of Flag- and His-tagged CSN5 was confirmed by IB using anti-Flag and anti-His antibodies, respectively (Left). IgG heavy chain (HC) and light chain (LC) are indicated.



**Fig. 2.** CSN5 MPN domain and specific extensions. (A) Conformational heterogeneity of the Ins-1 region in CSN5 (residues 97–131; brown), CSN6 (yellow; PDB code: 4E0Q), Mov34 (cyan; PDB code: 2O95), Prp8p (magenta; PDB code: 2OG4), AfJAMM (green; PDB code: 1O10), and SRU-2040 (purple; PDB code: 2KCO). Secondary structure elements around CSN5 Ins-1 are shown for context in white. (B) Sequence alignment focused on the residues 97–171 (CSN5 numbering) of four human MPN+JAMM representatives. Residues that correspond to the JAMM motif (with the exception of E76) are indicated with a black \*. In addition to each of the proteins, conservation of the residues across the species are illustrated by a color code ranging from red for strictly conserved residues to light gray for residues that are highly variable.

unbound form or in complex with its K63-Ub2 substrate (20). Therefore, it provides for this enzyme family a model for a catalytically competent active site and for substrate interactions. Comparison of the zinc-binding sites of CSN5 and AMSH-LP revealed that the overall topology of their active sites is well conserved (Fig. 3A; *SI Appendix, Table S5*). In addition, the position and environment of the Gly76-Lys63 isopeptide, inferred from the AMSH-LP/K63-Ub2 complex and straightforwardly placed in the CSN5 active site confirmed that CSN5 adopts a catalytically competent geometry (Fig. 3B). The Gly76-Lys63 isopeptide bond, placed in the CSN5 zinc-binding site, is maintained via a hydrogen bond between the Gly76 carbonyl group and the Ser148 hydroxyl group and between the Lys63  $\epsilon$ -amino group and the Glu76 carboxylate. Moreover, analysis of the crystal packing showed that the C-terminal segment (residues 248–257) of a monomer docks into the active site of the neighboring molecule (*SI Appendix, Fig. S4C*). Interestingly, a close inspection revealed that this C-terminal portion interacts with the neighboring molecule's active site in a manner that is reminiscent of the K63-Ub2 isopeptide bond in the AMSH-LP/K63-Ub2 structure (Fig. 3B). Moreover, the CSN5 Ser254 hydroxyl group is hydrogen bonded to Ser148 and occupies a position similar to the distal ubiquitin Gly76 carbonyl in the AMSH-LP active site.

We also investigated the role played by the catalytic zinc ion on the structure and stability of the active site. The side-chain motions of the zinc catalytic site amino acids were analyzed. Their positions were stable over the course of the MD simulations (*SI Appendix, Fig. S4D*), and their averaged interatomic distances from  $\text{Zn}^{2+}$  were in good agreement with those measured from the CSN5<sub>1–257</sub> and AMSH-LP crystal structures (*SI Appendix, Table S5*).

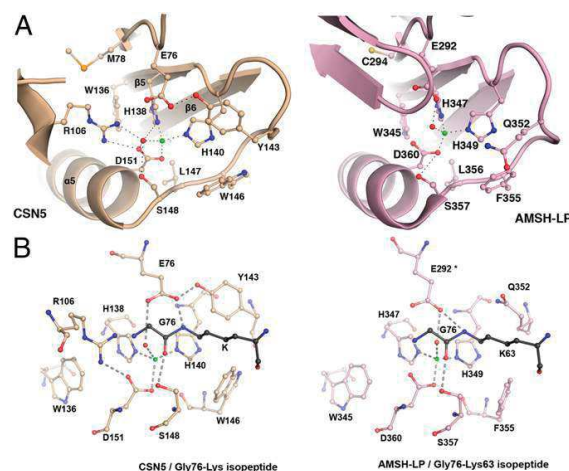
Taken together, these observations suggest that, as in AMSH-LP, the zinc-binding site catalytic residues of CSN5 are in a position and are geometry compatible with isopeptidase activity, and therefore the zinc active site conformation of this enzyme in its isolated form is catalytically competent.

Although the CSN5 zinc-binding site and its catalytic residues are very similar to those of AMSH-LP, their spatial environment and accessibility have several significant differentiating features.

In particular, the CSN5 Ins-1 region shielding the active site adopts a radically different topology (loop  $\beta$ 4- $\alpha$ 4 and  $\alpha$ 4 helix) compared with AMSH-LP characterized by two antiparallel  $\beta$ -strands followed by a short  $\alpha$ -helix (Fig. 4A). An additional crucial feature of CSN5 is the presence of the Arg106 residue projecting out of the Ins-1 segment and establishing a salt bridge with the Asp151 (Fig. 3A).

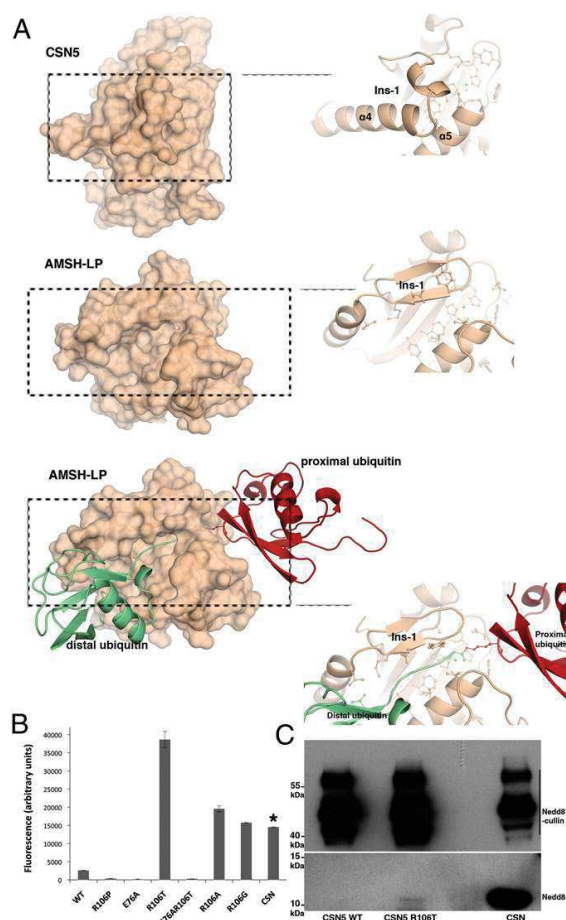
**Surroundings of the CSN5 Zinc Catalytic Site Is Not Competent for Nedd8 Recruitment, Without Conformational Rearrangements.** Two different activation states of CSN5 are described in the literature (4, 5): an active deneddylase in the context of the holo-CSN complex and a stand-alone inactive form in the isolated subunit. Because our data suggest that the CSN5 active site is poised for catalysis, it seemed logical to explore substrate binding and recruitment by this enzyme.

In the crystal structure of the AMSH-LP/K63-Ub2 complex, the two ubiquitin molecules, referred to as proximal and distal, interact with AMSH-LP via numerous electrostatic and hydrophobic interactions (*SI Appendix, Fig. S2*) (20). The directionality of the isopeptide bond implies that neural precursor cell expressed developmentally downregulated gene 8 (Nedd8) would occupy the site corresponding to the distal ubiquitin in the AMSH-LP/K63-Ub2 structure. The distal ubiquitin molecule mediates the largest interaction surface area and contributes the most to the binding affinity of K63-Ub2 for AMSH-LP. Correct positioning of the K63-Ub2 isopeptide bond in the long recognition groove of AMSH-LP is ensured by interactions between AMSH-LP [in particular, the Ins-1 region, the Ins-2 loop (disordered in CSN5), and the segment between these two insertions] and the proximal and distal ubiquitins (*SI Appendix, Fig. S2*). The C-terminal portion of the distal ubiquitin adopts an extended conformation that fits in the substrate-binding groove delimited by two  $\alpha$ -helices and a  $\beta$ -hairpin. Ubiquitin and Nedd8 molecules are 58%



**Fig. 3.** Geometry of CSN5 zinc-binding site is compatible with catalysis. Zinc and catalytic water are represented by green and red spheres, respectively. (A) CSN5 active site residues are poised for catalysis. Active site residues and their associated secondary structure are presented for CSN5 (light brown) and AMSH-LP (pink). (B) Model of AMSH-LP Gly-Lys isopeptide in CSN5 confirms that its zinc-binding site geometry is similar to that of AMSH-LP and is therefore competent for isopeptidase activity. The residues of the CSN5 and AMSH-LP zinc-binding sites and of the Gly-Lys isopeptide are shown in ball-and-stick format (light brown, light pink, and dark gray, respectively). The positions of the AMSH-LP residue Glu292, marked by an asterisk, as well as its zinc and catalytic water, were extrapolated from the AMSH-LP structure (PDB code: 2ZNR); the rest is from the structure of the AMSH-LP/K63-Ub2 complex (PDB code: 2ZNV).





**Fig. 4.** CSN5 adopts a conformation incompatible with Nedd8 recruitment and requires conformational relaxation to perform catalysis. (A) CSN5 Ins-1 masks the putative binding site for Nedd8. (Top) CSN5 surface and a close-up view of the Ins-1 region. (Middle) AMSH-LP surface and a close-up view of the Ins-1 region. (Bottom) Surface of AMSH-LP bound to K63-Ub2 and a close-up view of the groove that accommodates the isopeptide. (B) R106T (or R106A,G) substitution constitutively activates CSN5 against Nedd8 substrate. Isopeptidase activity of CSN5 or of the CSN on Nedd8-AMC was measured by the increase of fluorescence intensity at 460 nm ( $\lambda_{\text{excitation}} = 380$  nm) after 50 min at 28 °C. The concentration of CSN5 used in these experiments is 210-fold more than that of the CSN complex, as indicated by an asterisk. Taking into account this dilution factor, the CSN complex appears 79-fold more active than the most active R106 variant, R106T. (C) The R106T variant form of CSN5<sub>1–257</sub> is able to deneddylate Nedd8-cullin 1. (Upper) Nedd8 signal of the Nedd8-cullin 1 and (Lower) Nedd8 released on Nedd8-cullin 1 isopeptide bond hydrolysis for CSN5 WT, CSN5 R106T, and the CSN complex (from left to right).

identical over 76 residues and adopt the same fold (22, 23). The interactions with the last four residues of ubiquitin/Nedd8 preceding the isopeptide bond are likely to be preserved in CSN5. Only one residue at position 72 (arginine and alanine in ubiquitin and Nedd8, respectively) differentiates ubiquitin from Nedd8 in the last 10 residues. Analysis of the AMSH-LP residues implicated in the distal ubiquitin recognition site revealed that more than 65% of them are highly conserved in CSN5 (SI Appendix, Fig. S2). However, most of the residues for which no equivalent could be found in CSN5 belong to the highly divergent Ins-1 region that

has a very different sequence (Fig. 2B) and conformation in CSN5 and AMSH-LP. Consequently, without the structure of CSN5 in its active state, detailed analysis of the substrate-binding site in CSN5 is prevented.

Despite the high conservation of the interaction site in CSN5 (SI Appendix, Fig. S2), the conformation of the Ins-1 segment observed here sterically precludes Nedd8 binding (Fig. 4A). Extensive structural changes of this segment, which probably confers some of the specificity for Nedd8 ligand, would be required to create a fully competent binding site.

**Arginine Residue Contributes to the Control of CSN5 Isopeptidase Activation State.** The major difference at the active site level between CSN5 and AMSH-LP corresponds to the conformation of the Ins-1 segment. It is therefore most interesting to note that the Ins-1 segment of CSN5 shows signs of flexibility, as indicated by high crystallographic *B*-factor values and the fact that it exhibits significant conformational variability within representatives of the MPN family (Fig. 2A). Moreover, MD simulations aggregated portions of the Ins-1 region as highly flexible (residues 101–104, 108–112, and 123–131), with a propensity to adopt different conformations during the course of the simulations (SI Appendix, Fig. S4B). These CSN5 segments, bracketing the residue Arg106, display pronounced movements opening onto the solvent in MD simulations, whereas Arg106 contributes significantly to the anchoring of the Ins-1 segment to the zinc-binding site via its salt bridge with Asp151 (SI Appendix, Fig. S5A). Our MD studies confirmed the potential importance of Arg106 with this salt bridge being maintained in the 40-ns trajectory. The observations that Arg106 plays a role in CSN5 plasticity were further probed and confirmed by rotamerically induced perturbation (RIP) simulations (SI Appendix, Fig. S5B) (24). These data suggest that the intrinsic flexibility and plasticity of the Ins-1 region could allow major conformational rearrangements to accommodate Nedd8 binding and that Arg106 could have here a triggering function for structural rearrangement of the Ins-1 segment.

To evaluate the role of Arg106 as a potentially important protagonist in CSN5 activation switch, we tested the effect of Arg106 substitution on CSN5 isopeptidase activity. In agreement with published data in the literature (4, 5), we confirmed that the CSN5<sub>1–257</sub> WT form is void of isopeptidase activity and showed that the R106 substitution to threonine is sufficient to restore robust constitutive isopeptidase activity against two synthetic substrates: LRGG-7-amido-4-methylcoumarin (AMC) and Nedd8-AMC (Fig. 4B; SI Appendix, Fig. S6). Substitution of the arginine residue by alanine and glycine similarly confers isopeptidase activity to CSN5, albeit to a slightly reduced level compared with the R106T variant, whereas the R106P variant is void of activity (Fig. 4B). Analysis of the zinc binding site topology in CSN5 and AMSH-LP revealed that the glutamate residue corresponding to the position 292 in AMSH-LP is the glutamate 76 in CSN5. The position 76 was consequently mutated to an alanine residue. As expected and in comparison with the WT and R106T variants, the E76A and E76A/R106T variants do not display isopeptidase activity on Nedd8-AMC (Fig. 4B). Comparison of the rescued activity of CSN5 R106T with that of the holo-CSN complex using Nedd8-AMC as a substrate highlighted that the multiprotein complex is about 79-fold more active than the stand-alone protein (Fig. 4B). To complement these activity data, pull-down experiments, using GST-CSN5<sub>1–257</sub> as the bait and Nedd8 as the target, showed that the R106T form was much more efficient at binding Nedd8 than the WT counterpart (SI Appendix, Fig. S7). Furthermore, MD simulations on the R106T variant confirm that the Ins-1 segment is released from the vicinity of the zinc-binding site; similar results were obtained for R106G and R106P (SI Appendix, Figs. S4B and S5C). These observations confirm that releasing the Ins-1 segment from its anchoring point could favor a conformation that is compatible with Nedd8 binding. Taken together, these data strongly suggest the implication of Arg106 in the active/inactive switch of CSN5. The importance of R106 is

further supported by its strict conservation in CSN5 across eukaryote species (Fig. 2B; *SI Appendix*, Fig. S2). These results would indicate that the conformational relaxation of the Ins-1 region allows substrate binding and additionally corroborates our analysis on the intrinsic topological competence of the zinc binding site for catalysis but also point out that the CSN complex brings further activation and/or activity determinants that makes CSN-associated CSN5 more active than its stand-alone form.

To extend our analysis to a physiological substrate, we asked whether R106T also supports deneddylation of neddylated cullin 1, in complex with Rbx1 *in vitro*. As expected, the CSN complex is very effective at deneddyating cullin 1, as visualized by the important decrease of neddylated forms of cullin 1 and concomitant release of free forms of Nedd8, detected with Nedd8 antibodies. Importantly, although the stand-alone CSN5 WT was unable to deneddylate cullin 1, a modest but significant level of free Nedd8 was released when using the R106T variant (Fig. 4C). Confirming our results using Nedd8-AMC, these observations show that the WT-isolated form of CSN5 is indeed void of deneddylase activity, that the loss of the R106-D151 salt bridge contributes to unveiling its deneddylase activity, and that the integral CSN is much more efficient at deneddyating cullin 1.

## Discussion

Mediated most likely through its deneddylase activity, the function of the CSN complex is important for cellular homeostasis, as highlighted by its implication in proliferative diseases (25). The sequence alignment of the CSN catalytic subunit, CSN5, from different organisms reveals highly conserved features throughout the sequence in agreement with its catalytic function within the CSN complex. The distribution of CSN5 between the holo-CSN, different sub-CSN complexes, and CSN-independent forms observed in a number of studies could play a role in CSN5 regulation, but clearly the critical and best understood function resides in the deneddylation of cullins.

Our work reveals that CSN5 can be found in different oligomeric states *in vitro* and in cell lysates and may predominantly follow a monomer-dimer equilibrium. The interaction between CSN5 and various partners has been investigated in previous studies, but only in its monomeric form (2). Its assembly in dimers could be relevant in mediating protein-protein interactions and subcellular localization of CSN5.

The main aspect in CSN5 biology that is addressed in this work is its activation state in the CSN-independent context. To glean insights into CSN5 isopeptidase activity regulation, we used structural biology and *in silico* MD simulations, which together created a detailed picture of CSN5 activity control. The crystal structure of CSN5 in a CSN-independent form displays an extended catalytic domain that revealed a number of features, contributing to our understanding of the enzyme's activation and substrate recruitment. In analogy to the structure of AMSH-LP (20), the stand-alone form of CSN5 adopts a zinc-binding site geometry that appears compatible with isopeptidase activity and potentially with binding of the Gly76-Lys63 isopeptide, as extrapolated from the cocrystal structure of AMSH-LP/K63-Ub2 to the CSN5 zinc-binding site (Fig. 3B). Unlike AMSH-LP/K63-Ub2, however, investigation of the recruitment of Nedd8 by CSN5 revealed that the substrate-binding site is not formed in CSN5 and that the Ins-1 segment would require substantial structural rearrangement for Nedd8 to bind. These observations were confirmed by analysis of Ins-1 flexibility and plasticity by *in silico* simulations. Our work also helped understanding one of the molecular events that may trigger these conformational changes in CSN5. MD and RIP calculations pointed to a role for the strictly conserved Arg106 in keeping this segment in a conformation not competent for Nedd8 binding. The implication of this residue, validated by *in vitro* experiments, led to the confirmation that Arg106 is an important protagonist in CSN5 activation switch. Indeed, substitution of this residue by a threonine and to a lesser extent, alanine and glycine (but not proline), restores a constitutive isopeptidase activity against two synthetic substrates: LRGG-AMC and Nedd8-AMC

(Fig. 4B). This work was extended by testing isopeptidase activity of the R106T variant on a physiological substrate, neddylated cullin 1. The CSN5 R106T variant shows modest but significant isopeptidase activity, confirming results obtained on synthetic substrates (Fig. 4C). However, comparison of this activity with that of the CSN complex reveals, as observed for Nedd8-AMC, that the CSN complex is a more efficient deneddylase than CSN5 R106T. It is noteworthy that, although we observed robust isopeptidase activity on Nedd8-AMC by the CSN complex, this is in contrast with the observations reported by the Dubiel's group (26, 27), in which no activity was found. Even though the CSN isopeptidase activity measured in our experiments corresponds to a weaker level than the activity recorded on a physiological substrate reported in ref. 28, probably reflecting the higher efficiency of the CSN complex on the physiological substrate, this activity was readily measured in our fluorescence-based assay. Additionally Pan's group reported that the CSN complex was unable to process another linear substrate, pro-Nedd8 (29), using a gel-shift assay. These discrepancies, probably reflecting on the complexity of the studied system, would need further investigation to be resolved.

These observations relative to CSN5 activation shed some light onto different regulatory levels that keep CSN5 under control. CSN5 appears to require an activation step for which one of the switches is likely to correspond to the Arg106-Asp151 salt bridge: loss of this strong interaction promotes the existence of an open form of the enzyme that unveils deneddylase activity on both synthetic and physiological substrates. Coming in addition to the Ins-1 release is its restructuring to form the second ridge of the substrate-binding groove in which the isopeptide would insert. Integration of CSN5 into the CSN complex and the consequent protein-protein interactions with CSN subunits such as CSN6, as highlighted by nondenaturing MS experiments (5), are likely to play a part in both CSN5 activation and substrate recruitment. Given the body of evidence (30–33) that supports the involvement of multiple CSN subunits in substrate recruitment (e.g., the interaction between CSN2 and cullin 1), the observed difference in activity between the CSN complex and the stand-alone R106T CSN5 could be directly linked to the presence of substrate recruitment exosites remote from CSN5 catalytic site, achieving optimal activation and contributing to productive binding of Nedd8-cullin isopeptide bond. Supported by our work, CSN5 incorporation into the CSN complex probably does not lead to global structural reshaping of the enzyme. Instead, the structural changes are likely to be limited to the conserved Ins-1 segment (identified as malleable in our MD calculations), the Ins-2 region (disordered in the crystal), and possibly the C-terminal domain (residues 258–334) to prime the deneddylation molecule for catalysis. Integration of CSN5 in the CSN complex is probably providing the conformational energy necessary for the activation switch.

In addition to addressing CSN5 activity regulation, our work generates insights into the development of CSN5-specific inhibitors. Essentially, the C-terminal tail of one molecule docks in the active site of another one. This association could be exploited to build potent peptidomimetic antagonists that would be of possible interest in the context of breast cancers (2).

Taken together, our study suggests that CSN5 in its CSN-independent form is deficient in substrate recruitment and that a single residue contributes significantly to the activation switch. This discovery provides the framework for conducting biochemical and functional investigations to elucidate the regulatory mechanisms that control CSN5 function.

## Materials and Methods

**Construct Design.** For detailed construct design, cloning, mutagenesis, protein expression, and purification procedures, see *SI Appendix*.

**Crystallization, Data Collection, Structure Determination, and Molecular Dynamics Simulations.** Purified CSN5<sub>1–257</sub> formed crystals using the sitting drop vapor diffusion method. A dataset, collected at 2.6-Å resolution from a SeMet-labeled CSN5<sub>1–257</sub> crystal, was used to determine the structure using the SAD

method. The Se substructure was solved using the Charge Flipping method (34) and used for phasing with the PHASER program (35). Details are provided in *SI Appendix* and *SI Appendix, Table S1*. The A chain from the CSN5<sub>1-257</sub> crystal structure was used as the initial structure for MD simulations. Detailed procedures and references are provided in *SI Appendix*.

**Coimmunoprecipitation Experiments and Isopeptidase Assay.** Details regarding cell culture, reagents, transfection, and coimmunoprecipitation experiments are provided in *SI Appendix*. Different types of AMC-derived substrates, LRGG-AMC, and Nedd8-AMC, as well as neddylation cullin 1, were used to assess CSN5 activity. Details are provided in *SI Appendix*.

**ACKNOWLEDGMENTS.** We thank N. Zheng for constructive discussion and expression plasmids; M. Cohen-Gonsaud, J. Endicott, O. Coux, and L. Chiche for helpful discussions; and D. Staunton, L. Coquet, and A. Le Roy for experimental assistance. This work was supported by the Université Montpellier 1, the Institut National de la Santé et de la Recherche Médicale, the CNRS [Projets Exploratoires Premier Soutien (PEPS) Grant 2009], and the Groupement des Entreprises Françaises dans la lutte contre le Cancer (GEFLUC) Languedoc Roussillon (A.E.); the National Cancer Institute (RO1-CA90853) and Cancer Prevention and Research Institute of Texas (RP120451) (to F.X.C.); and the CNRS, the Association pour la Recherche sur le Cancer, and the City of Paris (L.P.). We acknowledge the use of the ID14-EH4 beamline [European Synchrotron Radiation Facility (ESRF)].

- Wei N, Serino G, Deng XW (2008) The COP9 signalosome: More than a protease. *Trends Biochem Sci* 33(12):592–600.
- Shackelford TJ, Claret FX (2010) JAB1/CSN5: A new player in cell cycle control and cancer. *Cell Div* 5:26.
- Claret FX, Hibi M, Dhut S, Toda T, Karin M (1996) A new group of conserved co-activators that increase the specificity of AP-1 transcription factors. *Nature* 383(6599):453–457.
- Cope GA, et al. (2002) Role of predicted metalloprotease motif of Jab1/CSN5 in cleavage of Nedd8 from Cul1. *Science* 298(5593):608–611.
- Sharon M, et al. (2009) Symmetrical modularity of the COP9 signalosome complex suggests its multifunctionality. *Structure* 17(1):31–40.
- Maytal-Kivity V, Reis N, Hofmann K, Glickman MH (2002) MPN+, a putative catalytic motif found in a subset of MPN domain proteins from eukaryotes and prokaryotes, is critical for Rpn11 function. *BMC Biochem* 3:28.
- Freilich S, et al. (1999) The COP9 signalosome is essential for development of *Drosophila melanogaster*. *Curr Biol* 9(20):1187–1190.
- Kwok SF, et al. (1998) Arabidopsis homologs of a c-Jun coactivator are present both in monomeric form and in the COP9 complex, and their abundance is differentially affected by the pleiotropic cop/det/fus mutations. *Plant Cell* 10(11):1779–1790.
- Mundt KE, Liu C, Carr AM (2002) Deletion mutants in COP9/signalosome subunits in fission yeast *Schizosaccharomyces pombe* display distinct phenotypes. *Mol Biol Cell* 13(2):493–502.
- Oron E, et al. (2002) COP9 signalosome subunits 4 and 5 regulate multiple pleiotropic pathways in *Drosophila melanogaster*. *Development* 129(19):4399–4409.
- Tomoda K, et al. (2002) The cytoplasmic shuttling and subsequent degradation of p27Kip1 mediated by Jab1/CSN5 and the COP9 signalosome complex. *J Biol Chem* 277(3):2302–2310.
- Fukumoto A, Tomoda K, Kubota M, Kato JY, Yoneda-Kato N (2005) Small Jab1-containing subcomplex is regulated in an anchorage- and cell cycle-dependent manner, which is abrogated by ras transformation. *FEBS Lett* 579(5):1047–1054.
- Yao T, Cohen RE (2002) A cryptic protease couples deubiquitination and degradation by the proteasome. *Nature* 419(6905):403–407.
- Kapelari B, et al. (2000) Electron microscopy and subunit-subunit interaction studies reveal a first architecture of COP9 signalosome. *J Mol Biol* 300(5):1169–1178.
- Serino G, et al. (1999) Arabidopsis cop8 and fus4 mutations define the same gene that encodes subunit 4 of the COP9 signalosome. *Plant Cell* 11(10):1967–1980.
- Sanches M, Alves BS, Zanchin NI, Guimarães BG (2007) The crystal structure of the human Mov34 MPN domain reveals a metal-free dimer. *J Mol Biol* 370(5):846–855.
- Zhang H, et al. (2012) The crystal structure of the MPN domain from the COP9 signalosome subunit CSN6. *FEBS Lett* 586(8):1147–1153.
- Ambroggio XI, Rees DC, Deshaies RJ (2004) JAMM: A metalloprotease-like zinc site in the proteasome and signalosome. *PLoS Biol* 2(1):E2.
- Pena V, Liu S, Bujnicki JM, Lüthmann R, Wahl MC (2007) Structure of a multipartite protein-protein interaction domain in splicing factor prp8 and its link to retinitis pigmentosa. *Mol Cell* 25(4):615–624.
- Sato Y, et al. (2008) Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains. *Nature* 455(7211):358–362.
- Tran HJ, Allen MD, Löwe J, Bycroft M (2003) Structure of the Jab1/MPN domain and its implications for proteasome function. *Biochemistry* 42(39):11460–11465.
- Whitby FG, Xia G, Pickart CM, Hill CP (1998) Crystal structure of the human ubiquitin-like protein NEDD8 and interactions with ubiquitin pathway enzymes. *J Biol Chem* 273(52):34983–34991.
- Ramage R, et al. (1994) Synthetic, structural and biological studies of the ubiquitin system: The total chemical synthesis of ubiquitin. *Biochem J* 299(Pt 1):151–158.
- Ho BK, Agard DA (2009) Probing the flexibility of large conformational changes in protein structures through local perturbations. *PLOS Comput Biol* 5(4):e1000343.
- Kato JY, Yoneda-Kato N (2009) Mammalian COP9 signalosome. *Genes Cells* 14(11):1209–1225.
- Schmaler T, Dubiel W (2010) Control of deneddylation by the COP9 signalosome. *Subcell Biochem* 54:57–68.
- Hannss R, Dubiel W (2011) COP9 signalosome function in the DDR. *FEBS Lett* 585(18):2845–2852.
- Emberley ED, Mosadeghi R, Deshaies RJ (2012) Deconjugation of Nedd8 from Cul1 is directly regulated by Skp1-F-box and substrate, and the COP9 signalosome inhibits deneddylation of SCF by a noncatalytic mechanism. *J Biol Chem* 287(35):29679–29689.
- Wu K, et al. (2003) DEN1 is a dual function protease capable of processing the C terminus of Nedd8 and deconjugating hyper-neddylation CUL1. *J Biol Chem* 278(31):28882–28891.
- Huang X, et al. (2005) Consequences of COP9 signalosome and 26S proteasome interaction. *FEBS J* 272(15):3909–3917.
- Min KW, et al. (2005) CAND1 enhances deneddylation of CUL1 by COP9 signalosome. *Biochem Biophys Res Commun* 334(3):867–874.
- Olma MH, et al. (2009) An interaction network of the mammalian COP9 signalosome identifies Dda1 as a core subunit of multiple Cul4-based E3 ligases. *J Cell Sci* 122(Pt 7):1035–1044.
- Yang X, et al. (2002) The COP9 signalosome inhibits p27(kip1) degradation and impedes G1-S phase progression via deneddylation of SCF Cul1. *Curr Biol* 12(8):667–672.
- Dumas C, van der Lee A (2008) Macromolecular structure solution by charge flipping. *Acta Crystallogr D Biol Crystallogr* D64(Pt 8):864–873.
- McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.



## **Paper 2**

**Structure and Function of MPN (Mpr1/Pad1 N-Terminal) Domain-Containing Proteins**

## Structure and Function of MPN (Mpr1/Pad1 N-terminal) Domain-Containing Proteins

Melissa Birol and Aude Echali  r\*

Centre de Biochimie Structurale - UMR1054 - CNRS - INSERM - UMI, 29 rue de Navacelles, 34090 Montpellier Cedex, France

**Abstract:** MPN (Mpr1/Pad1 N-terminal) domain-containing proteins are present throughout all domains of life. In eukaryotes, MPN domain-containing proteins are commonly found in association with other molecules in large protein complexes, where examples comprise; the 26S proteasome and the COP9 (Constitutive photomorphogenesis 9) signalosome complexes, including the MPN subunits, POH1 and Mov34, CSN5 and CSN6, respectively. Examples of MPN domain-containing proteins that are not incorporated in a large multi-protein complex have also been reported and include AMSH (for associated molecule with the SH3 domain of STAM) and the AMSH-Like Protein (AMSH-LP). Within the MPN domain super-family, two main subclasses have been characterised: the MPN<sup>+</sup> and MPN<sup>-</sup> domain-containing proteins. MPN<sup>+</sup> domain-containing proteins are classified as metalloenzymes responsible for isopeptidase activity. These proteins display a JAMM (JAB1-MPN-MOV34) metalloisopeptidase motif, typically consisting of a canonical sequence (E-x[2]-H-S/T-H-x[7]-S-x[2]-D) and coordinating a zinc ion. The JAMM motif specifies a catalytic centre essential for selective hydrolysis of linkages, contained between ubiquitin/ubiquitin-like proteins and target proteins or between ubiquitin monomers within a polymeric chain. The MPN<sup>-</sup> family classifies proteins, which lack the key residues present in the typical JAMM motif. These MPN<sup>-</sup> proteins are void of catalytic activity, but recent studies have proposed a role in mediating protein-protein interactions, in acting as a scaffold or in activity regulation. In light of recent structural and functional studies, a more detailed understanding of these proteins has been gained and is given in the present review.

**Keywords:** Cell signalling, ubiquitin-proteasome system, isopeptidase activity, COP9 signalosome, Prp8, protein-protein interactions, ubiquitin, Nedd8, CSN.

### 1. INTRODUCTION

The MPN (Mpr1/Pad1 N-terminal) domain is ubiquitously found in prokaryotes and eukaryotes. Through bioinformatics analysis, fourteen MPN domain-containing proteins have been found in human [1]. Whereas the first MPN protein structure came from a prokaryotic organism [2, 3], most of the remaining structural and functional work on MPN domain-containing proteins has been carried out on eukaryotes (for structural studies, [4-14]). This is the reason why the present review mainly focuses on eukaryotic members. Most of these MPN domains are, either part of a longer polypeptide chain [15, 16], inserted in a multi-protein complex [17-21] or both [18]. The proteins that contain an MPN domain either display an isopeptidase activity (such as POH1) or are void of any catalytic activity (such as Mov34). Moreover, MPN domains present in multi-protein complexes are often found in pairs, forming heterodimers. Most studied MPN domain-containing complexes include the proteasome lid, the COP9 signalosome, BRCC36 (for BRCA1/BRCA2-containing complex subunit 36)-containing complexes and eIF3 (for eukaryotic translation initiation factor 3) [17, 18, 22].

Whilst the function of each MPN domain-containing proteins is not necessarily always well understood, MPN members are often directly linked to the ubiquitin-proteasome pathway [16, 17, 23-29]. Indeed, out of the 14 known members, 12 MPN domain-containing proteins are either part of a complex carrying an isopeptidase activity, (Abraxas, Abro1, BRCC36, CSN5, CSN6, eIF3f (for eukaryotic translation initiation factor 3 subunit f), eIF3h (for eukaryotic translation initiation factor 3 subunit h), Mov34, MYSM1 (for Myb-like, SWIRM and MPN domain-containing protein 1), POH1) [16, 17, 24, 26, 27, 29], or themselves a standalone deubiquitinylase (AMSH (for Associated Molecule with the SH3 domain of STAM), AMSH-LP (for AMSH-Like Protein) [7, 26]. The substrate specificity of these enzymes and complexes contributes to deciphering which signalling pathways they belong to. Indeed, in the case of poly-ubiquitin chains, there is a strong link between the topology of the chains, i.e., which ubiquitin lysine residue is used to form the isopeptide bond with the following ubiquitin molecule, and the biological function [30]. Most studied are the K48-linked poly-ubiquitin chains that are associated with proteasomal degradation, but other poly-ubiquitin linkage types, such as K11, are also linked to proteolysis. K63-poly-ubiquitin chains are generally involved in other signalling pathways, such as DNA damage response and endocytosis. Interestingly, these ubiquitin-proteasome system (UPS)-linked entities containing an MPN domain are linked to different signalling pathways, such as protein degradation (POH1/Rpn11,

\*Address correspondence to this author at the Centre de Biochimie Structurale - UMR1054 - CNRS - INSERM - UMI, 29 rue de Navacelles, 34090 Montpellier cedex, France; Tel: 003(0)467417902; Fax: 0033(0)467417732; E-mail: [aude.echali  r-glazer@cbs.cnrs.fr](mailto:aude.echali  r-glazer@cbs.cnrs.fr)

Mov34/Rpn8, CSN5, CSN6) [24, 29], DNA damage control (Abraxas, Abro1, BRCC36) [18], endocytosis (AMSH, AMSH-LP) [7, 26], protein biosynthesis (eIF3f, eIF3h) [31] and transcriptional regulation (MYSM1) [16]. Out of the two remaining members, namely MPND (for MPN domain-containing protein) and Prp8, very little is known about the function of the former, whilst the role of Prp8 in mRNA maturation has been investigated in detail [32]. Furthermore, as these entities are implicated in various signalling pathways, the MPN-containing subunits are often associated with pathologies. This is the case, for example, for CSN5, CSN6, MYSM1, eIF3f, eIF3h that are involved in cancers [31, 33–35]. Moreover, there is a direct link between mutations on the MPN domain of Prp8 and the degenerative eye disease retinitis pigmentosa type 13 [12, 15].

Through the structural and biochemical information available, elements that dictate substrate recruitment, specificity and activity regulation have emerged [7, 10]. In this review, we will discuss structural and functional insights into MPN domain-containing proteins, gained mainly from crystallographic contributions. In particular, we present the structural elements of the MPN domain, elaborations on the fold and the sequence conservation of the proteins, as well as their association with themselves and with other proteins. Moreover, we discuss the regulation mechanisms of the isopeptidase activity carried by MPN domains.

## 2. MPN DOMAIN-CONTAINING PROTEINS AND COMPLEXES

### 2.1. MPN Domain-Containing Proteins in Higher Eukaryotes

Fourteen main MPN domain containing-proteins have been found in human, namely Abraxas, Abro1, AMSH, AMSH-LP, BRCC36, CSN5, CSN6, eIF3f, eIF3h, Mov34, MPND, MYSM1, POH1, Prp8 [1] and their general characteristics are presented in Table 1. Most of them have been studied in detail; whilst little is known for others.

Investigation of MPN domain occurrence in proteins revealed that it is a ubiquitous domain found in animals, plants, bacteria, archaea and viruses [36]. Its presence in prokaryotes and viruses has been surprising, until recently, when a system homologous to that of the ubiquitin in eukaryotes, based on the protein Pup1, was found in prokaryotes [37].

### 2.2. The JAMM/MPN<sup>+</sup> and the MPN<sup>-</sup> Domains

Members of the MPN domain-containing protein super-family can be segregated in two main categories: the catalytically active (JAMM/MPN<sup>+</sup>) and the catalytically void (MPN<sup>-</sup>) MPN domain-containing proteins. JAMM/MPN<sup>+</sup> domain-containing proteins are zinc-dependent isopeptidases and display a zinc-coordinating JAMM (JAB1-MPN-MOV34) metalloisopeptidase motif typically consisting of the conserved sequence (E-x[2]-H-S/T-H-x[7]-S-x[2]-D), as defined and validated by seminal mutational studies in *Saccharomyces cerevisiae* [24, 29, 38]. The JAMM motif specifies a catalytic centre essential for selective hydrolysis of isopeptide or peptide linkages contained between ubiquitin/ubiquitin-like proteins and target proteins or between ubiquitin monomers within a polymeric chain. The members

of the MPN<sup>-</sup> family that comprises proteins lacking the key residues present in the typical JAMM motif are bare of catalytic activity.

Phylogenetic analysis of the MPN fold super-family revealed that prokaryotic members are mostly JAMM/MPN<sup>+</sup> proteins [25, 38]. This observation suggests that MPN<sup>-</sup> members evolved from JAMM/MPN<sup>+</sup> proteins that lost the catalytic motif, but kept the JAMM/MPN<sup>+</sup> typical metalloprotease fold.

Whereas JAMM/MPN<sup>+</sup> and MPN<sup>-</sup> are distinct proteins, they are often found in pairs in multi-protein complexes. This is the case for the proteasome lid and the COP9 complex that have one representative of the JAMM/MPN<sup>+</sup> family, POH1 and CSN5 and one of the MPN<sup>-</sup> family members, Mov34 and CSN6, respectively [19, 20]. A similar JAMM/MPN<sup>+</sup>-MPN<sup>-</sup> composition is also found in the BRISC (for BRCC36 Isopeptidase Complex) and RAP80 complexes [17, 18]. The eukaryotic translation initiation factor eIF3 also contains one member of each MPN family, eIF3f (JAMM/MPN<sup>+</sup>) and eIF3h (MPN<sup>-</sup>) [21]. POH1, CSN5, BRCC36 and eIF3f correspond to the catalytic subunit of these complexes, but the MPN<sup>-</sup> specific role is still unclear.

### 2.3. Functions of MPN Domain-Containing Assemblies

The MPN domain-containing proteins are often found embedded in large multi-protein complexes. The composition and function of some of these multi-subunit entities are detailed below. A schematic representation of six of these assemblies is provided in (Fig. 1).

The **COP9 signalosome** is a multi-protein complex composed of eight subunits involved in the regulation of diverse cellular processes, including cell cycle progression, gene expression and DNA repair. The most studied activity of the COP9 complex corresponds to its deneddylation activity of cullin RING E3 ubiquitin ligases (CRLs) [39]. The catalytic activity of CRLs is controlled by the covalent attachment of an ubiquitin-like protein, Nedd8 (for Neural precursor cell expressed developmentally down-regulated protein 8), to a conserved lysine of the CRL cullin subunit via an isopeptide bond. The COP9 complex is responsible for the hydrolysis of this isopeptide bond and, thus, deneddylates cullins. This activity appears to be central to the CRL regulation. The isopeptidase activity of the COP9 complex is centred on its zinc-coordinating JAMM/MPN<sup>+</sup> domain-containing subunit, CSN5 [10, 24]. Although the MPN<sup>-</sup> subunit of the COP9 complex, CSN6, is thought to contribute to the structural integrity of the protein complex, this view was challenged by the atypical composition of yeast COP9 complexes. Indeed, the COP9 complex in yeast species diverges from the archetypical higher eukaryote one. In *Schizosaccharomyces pombe*, the COP9 complex is composed of six subunits, missing CSN6 and CSN8 [40]; in *Saccharomyces cerevisiae*, Csi1, a previously identified interactor of the COP9 complex essential for cullin deneddylation may be the CSN6 ortholog. Csi1 displays a significant sequence homology with the carboxy-terminal domain of CSN6, but not with its amino-terminal MPN domain [41]. It appears that the C-terminal domain of CSN6 alone could be sufficient for assembly, as well as deneddylating activity, suggesting that the MPN<sup>-</sup> domain could be accessory to the

Table 1. The human MPN super-family.

Proteins	Synonyms	Length (in amino acids)	Biological Pathway(s)	Catalytic Activity	Substrate Specificity/Role	Host Stable Complex
<b>Abraxas</b>	Abra1/CCDC98/FAM175A	409	DNA damage	MPN <sup>-</sup>	scaffold/activation	Rap80
<b>Abro1</b>	KIAA0157/FAM175B	415	endocytosis	MPN <sup>-</sup>	scaffold/activation	BRISC
<b>AMSH</b>	STAMBP	424	endocytosis	JAMM/ MPN <sup>+</sup>	K48- and K63-linked Ub chains	N.R.
<b>AMSH-LP</b>	STAMBPL1/KIAA1373	436	endocytosis	JAMM/ MPN <sup>+</sup>	K63-linked Ub chains	N.R.
<b>BRCC36</b>	C6.1/CXorf53	316	DNA damage/endocytosis	JAMM/ MPN <sup>+</sup>	K63-linked Ub chains	BRISC; Rap80
<b>CSN5</b>	COP5/JAB1	334	Protein degradation/COP9 function	JAMM/ MPN <sup>+</sup>	Nedd8-cullins	COP9
<b>CSN6</b>	COPS6/HVIP	327	Protein degradation/COP9 function	MPN <sup>-</sup>	N.R.	COP9
<b>eIF3f</b>	eIF3S5	357	Protein biosynthesis	JAMM/ MPN <sup>+</sup>	ubiquitinated protein substrate	eIF3
<b>eIF3h</b>	eIF3S3	352	Protein biosynthesis	MPN <sup>-</sup>	N.R.	eIF3
<b>Mov34</b>	Rpn8/PSMD7/S12	324	Protein degradation	MPN <sup>-</sup>	scaffold/activation	proteasome lid
<b>MPND</b>	/	471	unknown	JAMM/ MPN <sup>+</sup>	unknown	N.R.
<b>MYSM1</b>	KIAA1915	828	transcription regulation	JAMM/ MPN <sup>+</sup>	ubiquitinated histone 2A	p/CAF
<b>POH1</b>	Rpn11/PSMD14	310	Protein degradation	JAMM/ MPN <sup>+</sup>	ubiquitinated protein substrate	proteasome lid
<b>Prp8</b>	/	2335	mRNA maturation	MPN <sup>-</sup>	Br2 helicase stimulating role; ubiquitin binding	U5 snRNP

N.R.: Not reported in the literature.

COP9 activity in this biological context [41]. More broadly, the contribution of CSN6 within the COP9 complex remains to be established across different organisms.

One intriguing aspect of the COP9 complex is the fact that it is found in different sub-complexes in the cells [19, 42]. The physiological relevance and the function of these ‘mini-CSN’ complexes are not entirely clear. Moreover, the catalytic subunit, CSN5 is recurrently found outside of the COP9 complex, sometimes associated with other proteins, such as p27<sup>Kip1</sup> or c-Jun (reviewed in [33]). The link between these interactions and the ubiquitin-proteasome pathway has not yet been established.

**The proteasome lid** is part of the 19S regulatory particle (RP) of the proteasome, together with the base containing the AAA+ ATPases that form a hetero-hexameric ring (reviewed in [43]). The RP function is mainly linked to sub-

strate recognition, deubiquitinylation, unfolding and translocation of unfolded substrate to the proteasome core. More specifically, the lid JAMM/MPN<sup>+</sup> deubiquitinylase (DUB) activity carried out by the POH1 enzyme is to remove the entire poly-ubiquitin chain from the substrate protein at the attachment base and is essential for efficient substrate degradation [44]. Recently, the proteasome was shown to be able to cleave ubiquitin-Nedd8 mixed chains, increasing the range and the complexity of its known substrates; this activity was assigned to Rpn11, the POH1 ortholog in yeast [45]. Rpn11 is additionally involved in mitochondrial tubular organisation, cell cycle progression, transcriptional response to UV irradiation and also response to cellular damage, as investigated in yeast [46]. A second MPN domain-containing subunit, present in the proteasome lid is a catalytically inactive MPN<sup>-</sup> subunit, Mov34. Although the role of Mov34 remains to be precisely established, recent struc-

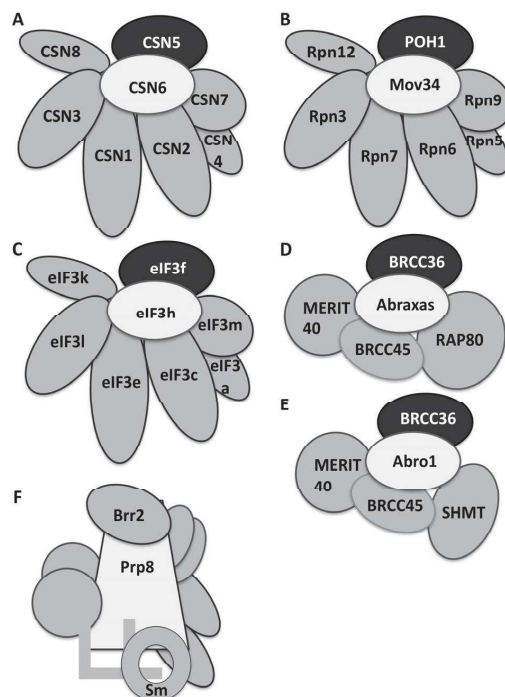
tural studies defined Mov34 as a direct partner of POH1 and as an important subunit for the proteasome lid assembly, integrity and activity [47-50].

**BRCC36-containing complexes.** BRCC36 is a JAMM/MPN<sup>+</sup> DUB that has been found in two complexes that have different sub-cellular distributions and functions, the RAP80 complex and the BRISC [51]. These complexes generally constitute structural cores that can subsequently associate with other elements, such as BCRA1 in the case of the RAP80 complex [17, 52] and the serine hydroxymethyltransferase (SHMT), for the BRISC. The five-member RAP80 complex consists of RAP80, BRCC36, BRCC45, Abraxas and MERIT40 (Fig. 1). The BRISC shares three subunits with the RAP80 complex, including BRCC36, BRCC45 and MERIT40 (Fig. 1). The fourth component of the BRISC includes the Abraxas paralog, Abro1, sharing 39% sequence identity with Abraxas [18]. In addition to the JAMM/MPN<sup>+</sup> BRCC36 subunit, these two complexes both comprise a MPN<sup>-</sup> member, Abraxas, for the RAP80 complex and Abro1, for the BRISC complex [17, 52]. Recently, a fifth component of the BRISC was identified, the SHMT that recruits the BRISC to the K63-ubiquitinated interferon receptor chain 1 (IFNAR1). Both of these complexes exhibit K63-ubiquitin-specific DUB activity [17]. The RAP80 complex is involved in DNA damage response (DDR) and associates with BRCA1 to form the BRCA1-A complex that is recruited at DNA damage sites. The newly found substrate, ubiquitinated IFNAR1, links the BRISC-SHMT assembly to the endocytosis machinery [51].

**AMSH and AMSH-LP** are standalone DUB proteins, which hydrolyse K48- and K63-linked polyubiquitin chains [26, 28]. They share 53% sequence identity and they both contain a nuclear localisation signal and a JAMM/MPN<sup>+</sup> domain. Furthermore, AMSH has been shown to bind clathrin and the Signal Transducing Adaptor Molecule (STAM), via its SH3-binding domain [28]. Although AMSH-LP possesses a SH3-binding-like motif, it is unable to associate with STAM [53], reflecting the fact that AMSH and AMSH-LP may have different cellular functions. The function of AMSH has been further probed and encompasses the regulation of receptor endocytosis. Specifically, AMSH contributes to the dynamics of receptor trafficking by deubiquitinating poly-ubiquitinated receptors and, therefore, by opposing E3 ubiquitin ligase activity [26]; AMSH-LP may carry out a related function. AMSH DUB activity, consequently, contributes to the regulation of free ubiquitin.

**Prp8** is a large polypeptide component of the U5 small nuclear ribonucleoprotein particle (snRNP; Fig. 1), which in combination with four other snRNPs, namely U1, U2, U4 and U6 and more than 100 proteins, forms the spliceosome. The spliceosome is a megadalton ribonucleoprotein particle, responsible for the removal of introns in eukaryotic pre-mRNAs and the re-ligation of the resulting RNA, in a mRNA maturation step referred to as splicing [32]. The U5 snRNP plays a critical role in the regulation of the spliceosome activity by interacting with a number of splicing factors, by affecting the spliceosome assembly, by contributing to the mRNA splice site selection and by contributing to the spliceosome active site formation in which the two transesterification reactions performed by the spliceosome take place. The MPN<sup>-</sup> domain of Prp8 located at its C-terminus

forms a complex with DEXD/H-box family helicase Brr2 and strongly stimulates its helicase activity [11, 13].



**Fig. (1).** Tentative schematic representation of multi-protein complexes that carry MPN domain-containing proteins. The following complexes are schematically represented in this figure: the COP9, the proteasome lid, the eIF3, two BRCC36-containing complexes, the U5 snRNP. The remaining MPN domain-containing entities (namely, AMSH, AMSH-LP, MPND, MYSM1) are not represented, as either they are stand-alone proteins, or they are not sufficiently characterised. These schematised complexes generally represent core components of larger assemblies responsible for the biological functions. JAMM/MPN<sup>+</sup> domain-containing proteins (CSN5; POH1; eIF3f; BRCC36) are represented in black; MPN<sup>-</sup> domain-containing proteins (CSN6; Mov34; eIF3h; Abraxas; Abro1; Prp8) in light grey. **Panel A:** the human COP9 complex that comprises eight subunits, among which two MPN domain-containing proteins (CSN5; CSN6) and 6 PCI-containing subunits. **Panel B:** the human proteasome lid with a core complex containing eight subunits, with two MPN (POH1; Mov34) and 6 PCI subunits. The proteasome lid is part of the regulatory particle of the 26S proteasome. **Panel C:** the human eIF3 core complex composed of eight subunits, with two MPN (eIF3f; eIF3h) and 6 PCI subunits. Another five subunits, namely eIF3b, -d, -g, -i, -j, complete this core assembly. **Panel D:** the core RAP80 complex is composed of five subunits, among which the pair of MPN domain-containing proteins, BRCC36 and Abraxas. **Panel E:** the BRISC complex comprises four subunits, among which the pair of MPN domain-containing proteins, BRCC36 and Abro1, plus a recently identified fifth subunit, SHMT. **Panel F:** the human U5 snRNP containing the U5 snRNA, Prp8, Brr2, the seven Sm proteins (B/B', D3, D2, D1, E, F, and G) and six other proteins (Snu114, Prp6, Prp28, 52K, 40K, Dibi).



**MYSM1** is a MPN domain-containing deubiquitylating enzyme [16, 54]. Cycles of histone ubiquitylation/deubiquitylation are an important regulatory process for transcriptional control. Among the many DUBs participating to this type of regulation, MYSM1 contributes via histone H2A deubiquitylation. It functionally possesses an intrinsic metalloprotease activity, resulting in the hydrolysis of the K63-linked ubiquitin chain. Although not entirely characterised and understood, MYSM1 appears to be part of a complex that may directly interact with the histone acetyltransferase (HAT) p300/CBP-associated factor (p/CAF) [16, 54]. The deubiquitylase activity of MYSM1 contributes to the cross-talk between different post-translational modifications, in particular, ubiquitylation and acetylation [16]. This MYSM1 activity modifies the pattern of proteins associated to nucleosomes. MYSM1 would therefore be a major player in transcriptional regulation. This is further confirmed in the context of hematopoiesis, in which MYSM1 appears to be an important protagonist, through its epigenetic function [35, 55]. The detailed molecular composition of the entities that include MYSM1 is not yet available.

**Human translation initiation factor 3 (eIF3)** is the largest eukaryotic initiation factor of the protein synthesis machinery. It is a multi-protein complex that carries out a central role in mRNA recruitment, as well as in the cellular translation machinery, where it contributes to the formation and assembly of translation initiation complexes. The eIF3 complex is composed of 13 non-identical subunits, designated eIF3a to -m. The subunits eIF3a, -c, -e, -k, -l, -m, -f and -h form an eight-subunit core (Fig. 1), on which the subunits eIF3d, -b, -g, -i and -j associate to form the full eIF3 assembly [56]. The eIF3f subunit of eIF3 is a JAMM/MPN<sup>+</sup> protein that exhibits deubiquitylase activity [27, 56]. As often for MPN<sup>+</sup> subunits, the eIF3 MPN<sup>+</sup> domain-containing protein, eIF3h, has a poorly defined function to date.

### 3. STRUCTURE OF MPN DOMAIN-CONTAINING PROTEINS

#### 3.1. Structural Data Available on MPN Domains

In recent years, considerable amount of structural information on MPN domain-containing proteins has become available, compiled in Table 2. Most of the known structures come from eukaryotic organisms, as only one prokaryotic MPN protein has been structurally and biochemically characterised [2, 3]. These two *Archaeoglobus fulgidus* AfJAMM structures provided the boundaries of a minimal MPN core domain and the first glimpse into the MPN domain structure (Fig. 2A, B). The structure of another prokaryote MPN, solved by NMR, has been reported in the Protein Data Bank (PDB; PDB code: 2KCQ), but the lack of associated publication on this protein restricts the analysis that can be done on it.

The structural data available comprises mostly MPN domain structures, determined, either alone or in the context of a larger polypeptide chain (Fig. 2A), with the exception of AMSH-LP (dead-mutant variant) that has been determined in complex with its substrate, a di-ubiquitin molecule (PDB code: 2ZNV; [7]) and of Prp8 that has been determined in complex with Brr2 (PDB codes: 4KIT; 4BGD) and with

Aar2 (PDB codes: 3ZEF; 4I43; 3SBT; 4ILG). Most recent structural information corresponds to the protein AMSH and its closely related AMSH-LP [7, 8], to subunits of the COP9 complex, namely CSN5 [10] and CSN6 [9], and to a large component of the spliceosome U5 snRNP, Prp8 [11-14]. Electronic microscopy studies have provided precious information, regarding the environment of these MPN proteins within their multi-protein complexes (reviewed in [43] for the proteasome; [16, 57-60]; for the COP9 complex, [58, 59]; for the eIF3 complex, [56, 60]). Taken together, these structural data provide solid grounds to understand the MPN domain architecture.

**Table 2.** Available structures containing MPN domains.

MPN Domain-Containing Protein	PDB Code	Boundaries of the Structure	References
AfJAMM	1R5X	1-119	[2]
	1OI0	1-119	[3]
HsAMSH-LP	2ZNR	264-436	[7]
HsAMSH-LP/K63-di-ubiquitin	2ZNV	265-436 E292A	[7]
HsAMSH	3RZU	248-424	[8]
HsCSN5	4F7O	1-257	[10]
DmCSN6	4E0Q	51-187	[9]
HsMov34	2O96	1-177	[5]
HsMov34	2O95	1-186	[5]
ScPrp8	2OG4	2148-2396	[4]
ScPrp8	3SBG	1836-2390	[57]
ScPrp8	3SBT	1834-2086	[57]
ScPrp8	4I43	885-2396	[11]
ScPrp8	4ILG	1836-2085; 2148-2392	[14]
CePrp8	2P87	2062-2307	[6]
ScPrp8	4BGD	2148-2395	[13]
ScPrp8	3ZEF	885-2387	[11]
HsPrp8	4KIT	2064-2325	[12]

Af: *Archaeoglobus fulgidus*; Hs: *Homo sapiens*; Dm: *Drosophila melanogaster*; Sc: *Saccharomyces cerevisiae*; Ce: *Caenorhabditis elegans*.

#### 3.2. Composition and Structure of MPN Domains

The structure of AfJAMM is organised in an almost closed eight-strand  $\beta$ -sheet that is completed with two  $\alpha$ -helices (Fig. 2B). Several regions could not be rebuilt from the electron density maps and are, therefore, missing in the final structure [2, 3]. The MPN core, derived from the seminal work on AfJAMM, contains an eight-strand  $\beta$ -sheet ( $\beta$ 1 to  $\beta$ 8) that is completed with three  $\alpha$ -helices. There are little variations in the secondary structure content of this MPN



core. The  $\beta 7$  and  $\beta 8$  strands, in some MPN structures, are fused and, in these cases, the structure contains seven strands (Mov34, Prp8; Fig. 2A, B). Despite the low sequence conservation between the MPN super-family members, superimposition of the different MPN domain-containing proteins, solved to date, confirms that this structural core (in beige) is conserved, and is completed by secondary structure elements (in green; Fig. 2B). Some loops are extended from the MPN core domain: this is the case between the strands  $\beta 2$ - $\beta 3$  (Prp8) and  $\beta 5$ - $\beta 6$  (CSN5). In addition, some secondary structure elements can vary in length: this is for example the case of the  $\alpha 3$  helix (reduction in Mov34 and in *AfJAMM*).

Further elaborating on the MPN fold, several eukaryotic MPN super-family members possess two insertions, referred to as Ins-1 and Ins-2, initially described in AMSH-LP structure [7] (Fig. 2A, B). AMSH-LP structure reveals the presence of these two insertions within the MPN fold, one between the  $\beta 3$  and  $\beta 4$  strands, referred to as Ins-1 and the second one, located at the C-terminus after the  $\beta 6$ , referred to as Ins-2. These insertions appear to hold considerable functional importance [7]. The AMSH-LP and AMSH Ins-1 segments form one ridge of the substrate-binding groove and the Ins-2 region contributes to the productive substrate positioning. The Ins-1 region appears to be present in most of the eukaryotic MPN domain-containing proteins examined, ranging from 24 and 35 residues. In the structure of CSN5, the Ins-1 conformation contributes to the activation state of this isopeptidase and, in AMSH-LP, to the substrate recruitment [7, 10]. This insertion greatly varies in terms of its sequence (from one member to another, but can be highly conserved for a given member across species [10]) and of its conformation (in AMSH-LP, CSN6 and Mov34, it contains a long helical portion and  $\beta$ -hairpin; in Prp8, the segment adopts an entirely extended conformation that may be stabilised by a Prp8-specific extension packing against the Ins-1 portion, completed by less than a helix turn; whilst the Ins-1 segment in CSN5 is mainly helical).

AMSH-LP and AMSH are the only two JAMM/MPN<sup>+</sup> DUBs characterised to date, in which Ins-2 conformation is constrained by a structural zinc. Whilst the Ins-2 segment contributes to substrate binding in AMSH-LP, as informed from its crystal structure in complex with its di-ubiquitin substrate, the role of this segment in other MPN super-family members is not clear. The size of the Ins-2 segment also greatly varies between members of the MPN super-family, with, in MPND, the Ins-2 being reduced to a 2-amino-acid linker between the antiparallel  $\beta$ -strands  $\beta 6$  and  $\beta 7$  and CSN5 Ins-2 segment being the longest (31 residues; disordered in the crystal structure; Fig. 2A, B). Interestingly, the Ins-2 segment in Prp8, although not constrained by a zinc, adopts a similar conformation to that of AMSH-LP and AMSH. The large variation in length, composition and conformation of the Ins-2 region raises the question of its function.

The crystal structure of the 51-187 fragment of *Drosophila melanogaster* CSN6 reveals a locally different topology of the C-terminal MPN fold compared to that of the other structurally-characterised MPN members determined to date [6]. Indeed, the  $\beta 5$  strand, although occupying the expected position in the structure, is oriented in the opposite direction compared to all the other MPN family members. It is possible that, because of the relatively short crystallised

CSN6 fragment, the  $\beta 5$  strand is enabled to adopt this atypical topology. Moreover, the Ins-2 segment of CSN6 is not encompassed in the crystallised fragment. Further structural work on CSN6 is needed to understand this point.

### 3.3. General Hydrolytic Activity Properties of JAMM/MPN<sup>+</sup> Proteins

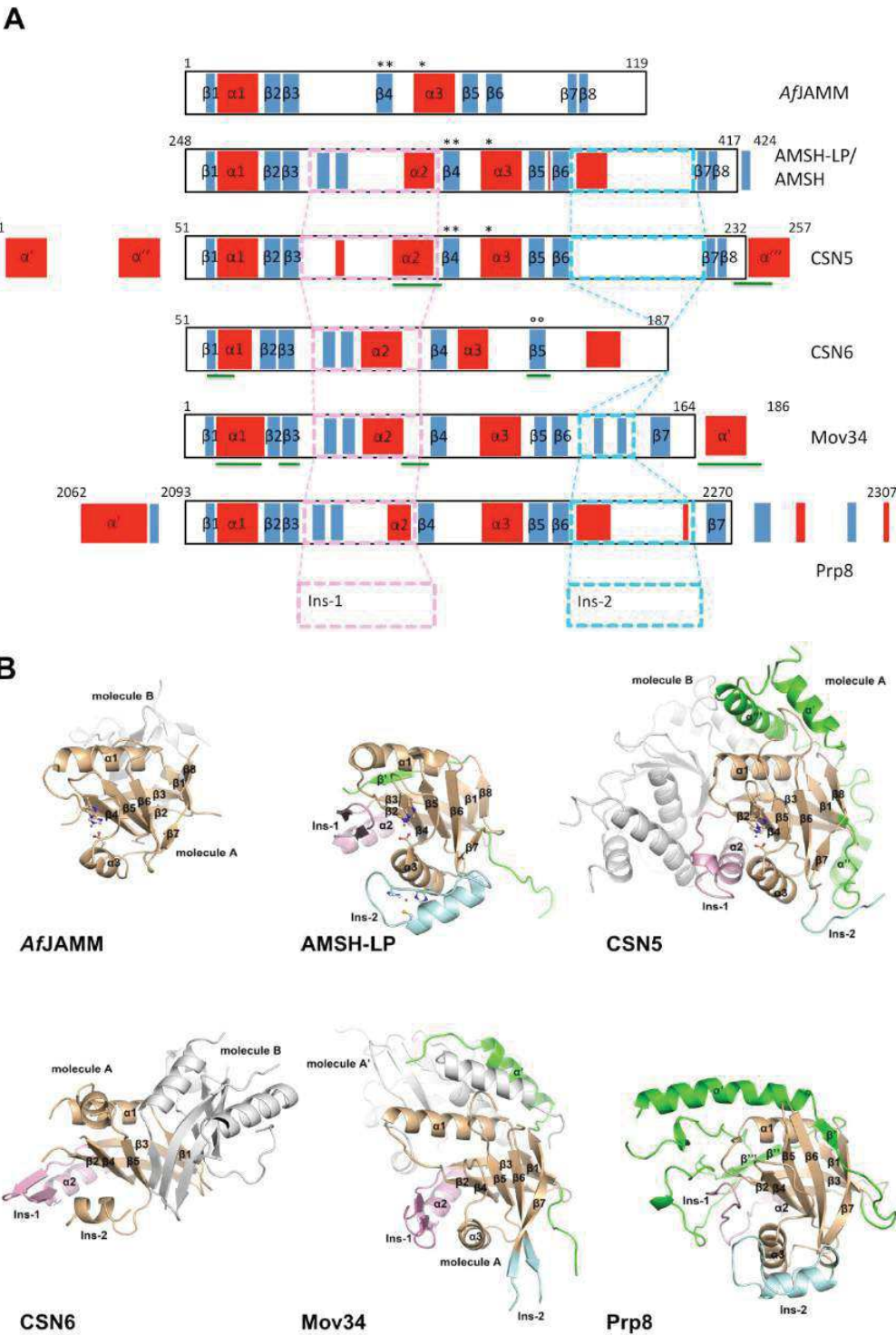
Among the JAMM/MPN<sup>+</sup> enzymes characterised to date, they are all isopeptidases, as reviewed and defined in [61, 62], with the exception of one prokaryote member, *Haloflex volcanii* JAMM1 that is also able to hydrolyse peptides [25]. Briefly, a functional active site comprising the zinc bound to the protein via two histidine and one aspartate residues, as identified in the proteasome lid [29,38] and the COP9 complex [24] and in the crystal structures of AMSH-LP [7], AMSH [8] and CSN5 [10] is necessary for their isopeptidase activity. The catalytic site is well conserved across the JAMM/MPN<sup>+</sup> family (Fig. 2B, C). It is noteworthy that, whilst eIF3f displays DUB activity, its catalytic site is difficult to align with that of other JAMM/MPN<sup>+</sup> members [27, 38]. As reviewed in [61, 62], the catalytic zinc is tetrahedrally coordinated to the polypeptide and to a water molecule. This catalytic water molecule is activated by the coordination to the zinc and its hydrogen bond to a conserved glutamate residue. Importance of this glutamate residue in catalysis was highlighted in AMSH-LP, where the E292A variant entirely loses its DUB activity [7] and was extended to CSN5 with the corresponding glutamate 76 residue [10]. Moreover, although detailed enzymatic study of the JAMM/MPN<sup>+</sup> activities is limited due to the complexity of some systems, kinetic characterisation of the COP9 complex was recently carried out [63], showing that unmodified CRLs could be potent COP9 inhibitors.

The active site of two JAMM/MPN<sup>+</sup> members, AMSH and CSN5, shares the same organisation and similar composition (Fig. 2C). Major peripheral differences in residue side chain composition include Q340 in AMSH that corresponds to Y143 in CSN5, F343 to W140 and T339 to G142, respectively, and could contribute to the substrate specificity and enzymatic characteristic. Even more striking is the presence of an arginine residue, R106 engaged in a salt bridge with the zinc-coordinating aspartate residue, D151 in CSN5. We investigated this significant difference between AMSH-LP and CSN5 active sites and showed that the R106-D151 salt bridge is involved in the activation state of CSN5 [10].

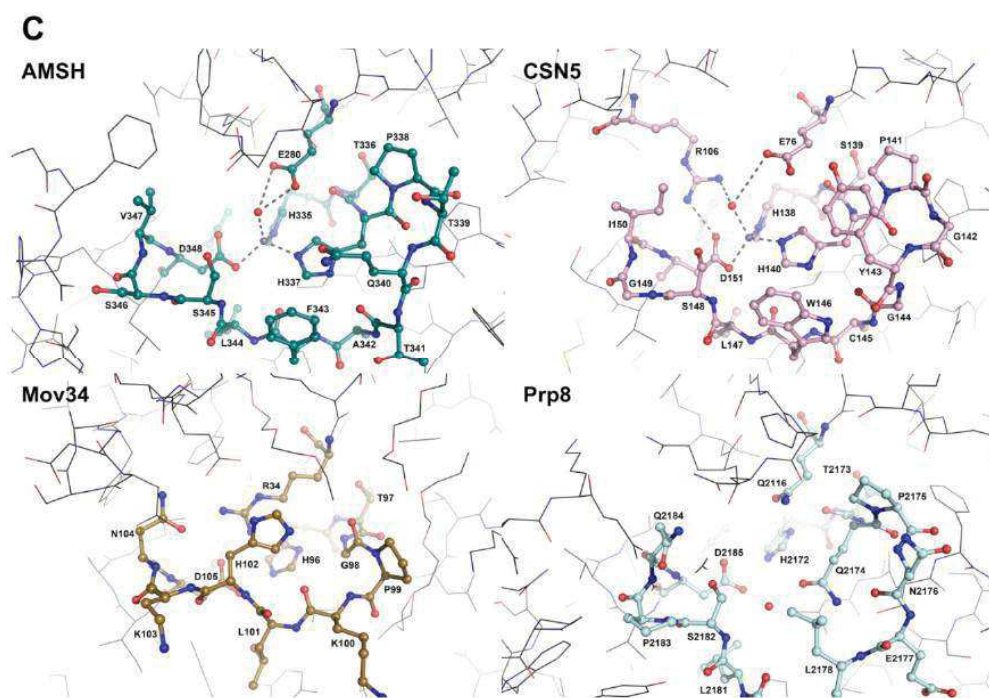
Comparison of the JAMM/MPN<sup>+</sup> zinc-binding site with the corresponding site in MPN<sup>-</sup> representatives reveals that a small vestigial cavity remains in the MPN<sup>-</sup> scaffolds (Fig. 2C). Moreover, although some striking differences between these JAMM/MPN<sup>+</sup> and MPN<sup>-</sup> sites exist, a few residues occupy equivalent positions, such as Q2116, H2172, S1282 and D2185, in Prp8, and H96 and D105, in Mov34, giving an overall similar topology to these sites. This is particularly marked in Prp8, which greatly resembles the sites of AMSH/AMSH-LP and CSN5.

### 3.4. Flexibility of the MPN Domain and Regions of Conformational Diversity

Whereas the minimal MPN core appears as a rigid domain through structural comparison between MPN



(Fig. 2) cond....



**Fig. (2). The MPN domain. Panel A:** Schematic representation of secondary structure elements of the MPN domain-containing proteins characterised to date. Boundaries of the MPN domain are delineated by a black rectangle. Secondary structure elements corresponding to MPN extensions are shown on both sides of the black box. Helices and strands are shown by red and blue boxes, respectively. The Ins-1 and Ins-2 segments are mapped by pink and blue dashed boxes, respectively. The positions of the zinc-coordinating residues for the three JAMM/MPN<sup>+</sup> members are indicated by asterisks. Green lines underneath CSN5, CSN6 and Mov34 show regions involved in crystal detected homo-dimerisation interfaces (A-B, A-B, A-A' dimers for CSN5, CSN6 and Mov34 1-186-fragment respectively). The  $\infty$  symbol indicates the CSN6  $\beta 5$  strand that adopts an atypical direction compared to the  $\beta 5$  strand of other MPN members. The analysis of the MPN core secondary structure content led to the identification of three helices  $\alpha 1$  to  $\alpha 3$  present in all MPN members, except for *A/JAMM* in which  $\alpha 2$  is not present. **Panel B:** Structure of the MPN domain-containing proteins highlighting the MPN core (in beige), its N- and C-terminal extensions (in green), the Ins-1 and Ins-2 regions (in pink and blue, respectively) and in white the most likely dimer molecule. **Panel C:** JAMM/MPN<sup>+</sup> active sites and MPN<sup>+</sup> equivalent sites. The active site residues are numbered and shown in ball-and-stick mode. Surrounding residues are shown in line mode for context. Interaction network around the zinc (light grey sphere) is shown with dash lines. The catalytic water is shown with a red sphere. Top left: AMSH (PDB code: 3RZU); Top right: CSN5 (PDB code: 4F7O); Bottom left: Mov34 (PDB code: 2O9S); Bottom right: Prp8 (PDB code: 2P87).

super-family members, B-factor value distributions and molecular dynamics simulations carried out on CSN5 [10], suggest that elements protruding from this MPN core display intrinsic conformational heterogeneity. These flexible structural elements include surface loops that have a high degree of freedom and segments that may be functionally plastic to mediate the protein function. Whilst it appears difficult to differentiate between these two without further functional and structural information, some regions appear to be more malleable across the different MPN members. Among these areas, the Ins-2 element recurrently shows signs of flexibility, either via high B-factor values (AMSH, AMSH-LP, Prp8), by being disordered (CSN5, Mov34) or by adopting different conformations in the MPN members (Fig. 2). Furthermore, portions of the Ins-1 segment show some signs of plasticity in AMSH-LP, CSN5 and CSN6. Our work on CSN5 activation showed that the plasticity of CSN5 Ins-1 segment is necessary for its activation [10] and

this observation could be relevant for other MPN domain-containing proteins.

### 3.5. N- and C-Terminal Extensions that Flank MPN Domains

Whereas, in prokaryotes, MPN domains that have been studied are mostly found in standalone polypeptides and rarely associated to accessory domains, in eukaryotes, the complexity of the MPN-containing entities is not limited to their MPN domain, but often extends to the presence of neighbouring domains within a long polypeptide chain. A striking example is the 2235-residues human protein Prp8, recently structurally characterised [11, 14], in which its MPN domain represents around 5% of the total polypeptide chain. Some members of the MPN super-family possess defined domains outside of the MPN domain. For example, MYSM1 and Prp8 display different structurally defined domains, among which a SANT (for, Swi3, Ada2, N-Cor, and TFIIIB)



and a SWIRM (for, Swi3, Rsc8, and Moira) domains, for MYSM1 and a reverse transcription homology, a restriction endonuclease homology and a RNase H homology domains, for Prp8. However, many MPN domain-containing proteins have extensions that are not easily classified. Proteins that contain an MPN domain largely vary in size and in quaternary arrangements and the presence of an MPN domain does not appear to be correlated with that of other formally defined domains. Interestingly, several MPN family members, such as Abro1, Abraxas, BRCC36, CSN5, CSN6, POH1/RPN11, Mov34/Rpn8, have a predicted helical region that may be implicated in complex assembly [18, 49, 58], following their MPN domain.

Detailed information that is available on structural elements flanking the MPN core domain is restricted to the crystallised polypeptide fragment. Probably for ease of heterologous expression and of crystallisation, many MPN domain-containing proteins crystallised to date (with the recent exception of Prp8, for which a large proportion of the protein (1502-residue fragment) has been crystallised; Table 2) correspond to the MPN portions of the polypeptide chain. Some fragments are limited to the MPN core, as it is the case for CSN6 and *A/JAMM*. Other structures include extensions around the MPN core (AMSH-LP; AMSH; CSN5; Mov34). Globally, these extensions flanking the MPN core domain wrap around it, deeply modifying the surface properties of the proteins (Fig. 2A, B). Two examples of these extensions are found in CSN5 and Prp8. In the crystal structure of CSN5 that encompasses the residues 1 to 257, the first 50 residues contain two  $\alpha$ -helices ( $\alpha'$  and  $\alpha''$ , respectively) and the last 26 amino-acids contain one extended  $\alpha$ -helix ( $\alpha'''$ ). The helices  $\alpha'$  and  $\alpha'''$  are in close spatial proximity and cross over the extended loop between the strands  $\beta 5$  and  $\beta 6$  (Fig. 2A,B). This structural motif is held by a strong hydrophobic contribution from tyrosine and tryptophan residues and is not found in other MPN domain-containing proteins structurally characterised to date. The helix  $\alpha''$  is placed above the  $\beta 2$  strand. For Prp8, the portion containing the MPN core includes N- and C-terminal extensions that comprise one  $\alpha$ -helix towards the amino-terminus packing in a groove above the helix  $\alpha 1$  and one  $\beta$ -strand complementing the central  $\beta$ -sheet of its MPN core (Fig. 2A, B). At the C-terminus, these extensions form an  $\alpha$ -helix and a small  $\beta$ -strand and the rest of the polypeptide chain tightly packs against the MPN core, adopting a spiral structure. The structural module constituted by the MPN core and these extensions is inserted into the remaining of the Prp8 structure via a long linker [12]. Prp8 is not the only MPN member to be inserted in the rest of the protein system via its C-terminal region. As for Abro1, Abraxas and CSN6, the C-termini of both proteasome lid subunits, POH1 and Mov34 are predicted to form long helices, which could form a coiled-coil module [49]. This module for POH1 and Mov34 is thought to serve as a flexible anchoring appendix into the remaining PCI-domain subunits.

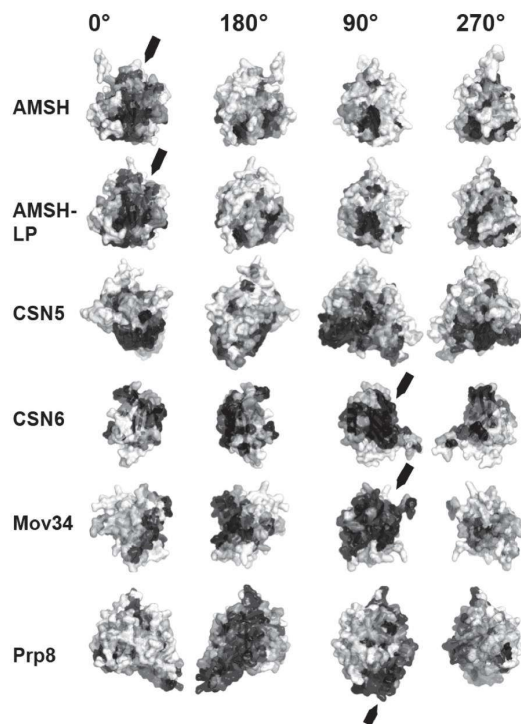
Overall, in the structures containing extensions of the MPN core domain, these appendices do not share structural composition, topology or organisation. It makes the MPN core an enzyme or protein-protein interaction platform tailored via specific extensions to carry out its specific func-

tion. Additional structural characterisation of integral MPN domain-containing proteins, as it has been done for Prp8, is therefore needed to fully grasp the properties of the molecules in their molecular context.

### 3.6. Conservation Surface Properties

The analysis of evolutionary highly conserved surfaces of proteins can be used as a diagnostic tool to determine functionally important regions and provide some clues on the role of these surfaces in the JAMM/MPN<sup>+</sup> and the MPN<sup>-</sup> family members. The surface conservation analysis on the MPN super-family members revealed several highly conserved patches (Fig. 3), as analysed by the program ConSurf [64].

As anticipated in JAMM/MPN<sup>+</sup> structures (AMSH, AMSH-LP and CSN5), the catalytic site cleft and the substrate-binding site groove, lined with the  $\alpha 2$  and  $\alpha 3$  helices, are very highly conserved, whilst the opposite face of the proteins is peppered with small variable patches (Fig. 3, 0°/180° views).



**Fig. (3). Conservation surface of MPN super-family representatives.** The conservation surface of six MPN super-family members calculated via the ConSurf server [64] is presented with a grey scale, in which black regions correspond to conserved residues and white to variable ones. Four orientations of the structures are presented: 0, 90, 180 and 270°, as indicated at the top of the panel. Arrows correspond to regions that are cited in the text.

Analysis of the conservation distribution among the MPN super-family members clusters JAMM/MPN<sup>+</sup> members, away from MPN<sup>-</sup> domain-containing proteins. In MPN<sup>-</sup>

members, the areas corresponding to the active site in the JAMM/MPN<sup>+</sup> members are not highly conserved, but a surface limited on one side by the  $\alpha 1$  helix and on the other side by the  $\alpha 2$  helix is strikingly preserved. Mov34 and CSN6 share this conservation pattern on their surface. The identified surface is likely to play a role in their functions. In CSN6 and Mov34, each side of the conserved surface is lined by a less conserved residue cluster. In the case of Prp8, this surface is slightly displaced towards the  $\alpha 1$  helix and reinforced by the  $\alpha'$ (Prp8) helix, whilst the  $\alpha 2$  helix is much reduced and highly variable.

#### 4. HIGHER ORDER ASSEMBLIES OF MPN DOMAIN-CONTAINING PROTEINS AND RELATIONSHIP WITH OTHER PROTEINS

##### 4.1. MPN Domain-Containing Proteins Form Homodimers in Crystal

Biochemical studies as well as the crystal structures of various MPN domain-containing proteins strongly suggest that some of them have the propensity to dimerise [9, 10, 19]. The assemblies are proposed to be potential homodimers according to their interface surface area, calculated by the PISA server [65] and to the additional biochemical or biophysical evidences (Table 3).

**Table 3. Summary of the molecule contacts in MPN domain crystal structures.** Interface surface areas were computed by the PISA server [65].

Protein	PDB Code	Surface Area ( $\text{\AA}^2$ )	Potential Homo-Dimer
<i>AfJAMM</i>	1OI0	720	X (?)
	1R5X	730	
<i>AMSH</i>	3RZU	700	
	3RZV	615	
<i>AMSH-LP</i>	2ZNR	550	
	2ZNV	260	
<i>CSN5</i>	4F7O	1055 974	X
<i>CSN6</i>	4EQ0	965 694	X (*)
<i>Mov34</i>	2O95	2300 600	X (?)
	2O96	1100 750	X (?)
<i>Prp8</i>	2P87	600	

(?): No additional experiments have been reported to establish the oligomeric state. (\*): Although it has been shown in cellular context that the full-length protein does not homo-dimerise.

**AfJAMM:** Crystal contacts between *AfJAMM* molecules in the 1R5X and 1OI0 structures reach an average surface area of  $730 \text{ \AA}^2$ . Although in each of these structures the interface surface implicates the same regions, namely the  $\alpha 1$

helix,  $\beta 2$  and  $\beta 3$  strands, it appears unlikely that this organisation corresponds to a biologically relevant assembly. Nevertheless, it was proposed that *AfJAMM* forms homodimers, even though in solution it has a monomeric behaviour [2].

**AMSH:** Interface areas between *AMSH* molecules in the crystals are also limited to around  $700 \text{ \AA}^2$ . The regions that are engaged in these interfaces mainly comprise the Ins-1, Ins-2 segments and the three last  $\beta$ -strands, among which  $\beta 7$  and  $\beta 8$ .

**AMSH-LP:** There is one molecule per asymmetric unit in *AMSH-LP* structure. Generation of symmetry-related molecules highlights only limited interface surfaces between molecules (at most to  $550 \text{ \AA}^2$ ) and no dimeric assemblies. The Ins-1 and Ins-2 segments are predominantly engaged in *AMSH-LP* interactions. In the *AMSH-LP*/di-ubiquitin structure, there is only very limited direct interactions between *AMSH-LP* molecules with a maximum interface surface of  $260 \text{ \AA}^2$ .

**CSN5:** Human CSN5 fragment 1-257 is organised in the crystal in two types of dimers, one contained in the asymmetric unit (A-B dimer) and the other generated by crystallographic operators (A-A' dimer). These two dimer types have a comparable interface surface, respectively  $1050$  and  $950 \text{ \AA}^2$ . Further investigation in solution is consistent with the preponderance of the asymmetric unit dimer (A-B) that is mainly maintained via interactions between the Ins-1 segment and the  $\alpha 4$  helix. Our study also confirmed that dimerisation of MPN domain-containing proteins is present in *celulo* by co-immunoprecipitation experiments on mammalian cell extracts and that the A-B dimer is preponderant. An intriguing aspect of CSN5 assembly into dimers is whether this has any relevance in mediating protein-protein interactions or whether this may be able to regulate CSN5 incorporation into the COP9 complex.

**CSN6:** Concerning *Drosophila* CSN6, dimers of the 51-184 fragment were observed both in the crystal and in solution [9]. Crystallographic and non-crystallographic symmetry operators generate two types of dimers with an interface surface of respectively  $700$  and  $970 \text{ \AA}^2$ , and SAXS experiments suggested that the asymmetric unit dimer, with the highest interface surface (A-B dimer) is the preponderant one. The A-B dimer is built on interactions between the  $\beta 1$ ,  $\alpha 1$ ,  $\beta 3$  and  $\beta 5$  structural elements and forms a slightly flattened  $\beta$ -barrel decorated by seven  $\alpha$ -helices. The direction inversion of the  $\beta 5$  strand might have an impact on the way CSN6 dimerises. However, the biological importance of the dimerisation of MPN domain-containing protein has been challenged in [41], where this work suggested that *in vivo* the full length CSN6 from mouse does not homo-dimerise. Surprisingly, CSN6 MPN domain alone appears to homo-dimerise *in vivo*.

**Mov34:** Two different fragments of human Mov34 have been crystallised (fragments 1-177 and 1-186; [5]). In terms of secondary structure content, they differ by the C-terminus  $\alpha 4$  helix that is present in the 1-186 fragment (PDB code: 2O95) and not in the other (PDB code: 2O96). In the two structures, the two Mov34 dimers adopt the same topology. For the 1-186-fragment structure, the asymmetric unit dimer (A-B) has an interface surface of  $750 \text{ \AA}^2$  and corresponds to the dimer generated by crystallographic symmetry operators

for the fragment 1-177 (interface surface of 600 Å<sup>2</sup>). In the 1-186-fragment structure, another dimer (A-A') generated by crystallographic operators (interface surface of 2300 Å<sup>2</sup>) corresponds to the asymmetric unit dimer (interface surface of 1100 Å<sup>2</sup>) in the 1-177-fragment structure. The 1-186-fragment A-A' dimer with its extended interface surface is stabilised by a helix swap ( $\alpha$ 4) from one monomer to the other and the lack of this  $\alpha$ 4 helix in the 1-177 fragment accounts for the dimer interface differences. Mov34 A-A' dimer (1-186-fragment) is maintained via numerous interactions between the helices  $\alpha$ 1,  $\alpha$ 2 and  $\alpha$ 4. Considering solvation free energies for the interface formation, the A-A' dimer interface may suggest a biologically relevant dimerisation interface. This A-A' dimer appears consequently as the most likely one, even so there is no further evidence of Mov34 homo-dimerisation in solution or *in vivo*. This point warrants further investigation.

**Prp8:** The asymmetric unit of Prp8 MPN domain (PDB code: 2P87) contains one molecule and symmetric molecules share only limited interface surface, ranging from 240 to 620 Å<sup>2</sup>. The interaction interface comprises residues scattered on the N-terminal  $\alpha$ -helix ( $\alpha'$ Prp8), the loop between  $\beta$ 2 and  $\beta$ 3 strands, the Ins-1 turn of  $\alpha$ -helix, the  $\alpha$ 3 helix and the  $\beta$ 7 strand. However, the interactions with neighbouring molecules are predominantly clustered around the N-terminal  $\alpha$ -helix. As Prp8 MPN domain is included in a large polypeptide chain, the relevance of Prp8 MPN domain dimerisation is limited and this is consistent with the limited interactions between Prp8 MPN domain molecules that probably result from crystal packing. Intriguingly, the analysis of the interaction of Prp8 MPN domain with the rest of Prp8 protein, as recently published (PDB code: 4I43; [11]), reveals that the MPN domain N-terminal  $\alpha$ -helix ( $\alpha'$ Prp8) is also implicated in this association. This observation reinforces the view that the helix is involved in protein-protein interactions.

Comparison of the dimer interface surfaces of CSN5, CSN6 and Mov34 indicates that these assemblies proceed via different surfaces (Fig. 2A, B). Although there are only limited secondary structure elements that are commonly implicated in these dimers, the general ability of MPN domain-containing proteins to dimerise has raised the question as to whether the observed homo-dimeric interactions may provide the basis for understanding the heterodimer formation of MPN domain-containing proteins present in the context of large protein complexes. It is noteworthy, however, that other MPN domain-containing proteins that have been crystallised, including AMSH-LP and Prp8, lack evidence of dimerisation.

#### 4.2. Organisation in Heterodimers

Often present in pairs in multi-protein complexes, the MPN domain-containing proteins have been shown to be able to associate to form heterodimers. Insights into their dimerisation were attained using yeast-two hybrid assays [66, 67], proteomic studies [68], non-denaturing mass-spectrometry [19, 20, 21] and electron microscopy [47, 48, 50, 56, 58-60]. Five large protein complexes are known to contain both JAMM/MPN<sup>+</sup> and MPN<sup>-</sup> domain-containing subunits. These include the 19S proteasome lid, the COP9 complex, the eIF3, and the BRCC36-containing complexes, BRISC and RAP80. In the proteasome lid, the POH1 and

Mov34 MPN domain-containing proteins form a heterodimer. This is thought to be also the case for CSN5/CSN6 in the COP9 complex, for BRCC36/Abraxas in the RAP80 complex and for BRCC36/Abro1 in the BRISC complex. Little is known about the way in which MPN domain-containing proteins assemble.

In the proteasome lid, the two MPN domain-containing proteins POH1 and Mov34 were thought to dimerise through their N-terminal MPN domains [5]. This topology was used for the hetero-dimer POH1-Mov34 in cryo-electron microscopy atomic model reconstruction of the proteasome, using the homodimer A-A' organisation found for the 1-186-fragment Mov34 [47]. Combining electron microscopy data with analysis of Mov34 MPN domain crystal structure enabled the construction of an atomic model of the heterodimer [47]. This study suggested that the interface of MPN dimerisation between Mov34 and POH1 is built on the same interface as the A-A' homo-dimer of Mov34 1-186 fragment, described in the previous section, but complementary studies would be needed to confirm this topology that remains speculative [50].

In agreement with the structural organisation of the proteasome lid, the elucidation of the electron microscopy map of the COP9 and the eIF3 complexes has indicated that the JAMM/MPN<sup>+</sup> and MPN<sup>-</sup> subunits are also in close proximity [58-60]. The MPN domain-containing proteins of both of these complexes are also likely to form a heterodimer, by interacting through their MPN domains.

There is also evidence suggesting that the enzymatic activity of BRCC36 requires one or more protein interactions occurring within each complex. Reminiscent of observations made on CSN5 and POH1, the BRCC36 DUB enzyme is inactive in its stand-alone state. BRCC36 only expresses its isopeptidase activity in the presence of the MPN<sup>-</sup> protein, Abro1 for BRISC or in the combined presence of the MPN<sup>-</sup> protein, Abraxas and of BRCC45 for the RAP80 complex. This activation of BRCC36 through the binding of other protein partners is correlated with its ability to form heterodimers with these MPN<sup>-</sup> proteins [18]. Mutational studies on the MPN<sup>-</sup> proteins accompanying BRCC36 have indicated that these proteins affect BRCC36 DUB activity possibly by contributing to substrate recruitment [18]. Similar to these observations, POH1 DUB activity, on K63-linked poly-ubiquitin chains, is exacerbated in the presence of Mov34 [18]. Reports suggest that the association between JAMM/MPN<sup>+</sup>-MPN<sup>-</sup> proteins could be mediated via the MPN domain, as well as additional domains such as the C-terminal coiled-coil domain in Abraxas and in Abro1 [69, 70].

Because MPN proteins appear to crystallise in homodimer through different interfaces, it still remains unclear whether all hetero-dimers of MPN domain-containing proteins proceed via the same dimerisation interface. However, for the time being, the description of the JAMM/MPN<sup>+</sup>-MPN<sup>-</sup> protein dimers at the atomic level is lacking and further structural work is required to be able to understand the hetero-dimerisation process of MPN domain-containing proteins.

By the very nature of their environments, in large polypeptide chains or in multi-protein complexes, MPN domains



are often found to mediate protein-protein interactions. Few examples for CSN5 are given in [33]. Although, at the moment, there is little data on the relevance and function of these protein-protein interactions, this is no doubt a central aspect to consider in the biology of these systems.

## CONCLUSIONS

The MPN domain has been found as part of numerous proteins in higher eukaryotes, prokaryotes and archaea. The MPN domain super-family shares a structural core amongst its members, but the presence of a catalytic motif in some led to the definition of two MPN family, one termed JAMM/MPN<sup>+</sup> and that carries an enzymatic activity and the other termed MPN<sup>-</sup> that is void of catalytic activity. In eukaryotes, the JAMM/MPN<sup>+</sup> proteins characterised to date display an isopeptidase activity specific to the ubiquitin or ubiquitin-like molecules, functionally linking the UPS to different pathways. Several JAMM/MPN<sup>+</sup> domain-containing proteins, together with MPN<sup>-</sup> domain-containing proteins, generally in pair are integrated in larger complexes, including the proteasome lid, the COP9, the eIF3 and the BRCC36-containing complexes.

The structural characterisation of six eukaryote MPN domain-containing proteins in the last six years has significantly contributed to a better understanding of this super-family. More specifically, the MPN domain appears as a structural core on which extensions are grafted. Large insertions, namely the Ins-1 and Ins-2 as first described in [7], procure specific features to the different members, linked to substrate recruitment, catalytic activity and protein-protein interactions, as highlighted in the JAMM/MPN<sup>+</sup> protein AMSH-LP. The presence of N- and C-terminal extensions to the MPN core radically changes the surface properties of these proteins, as shown in the structures of Prp8 and of CSN5.

Another specific characteristic of some MPN domain-containing proteins is their propensity to associate in homo-/hetero-dimers and to be implicated in protein-protein interactions. The analysis of crystal content revealed different homo-dimeric topologies and it is yet unclear if any of these arrangements are relevant to the way MPN domain-containing proteins hetero-dimerise *in vivo*. These proteins are often found in JAMM/MPN<sup>+</sup>-MPN<sup>-</sup> pairs and it has been shown for some that the binding of the MPN<sup>-</sup> molecule activates the JAMM/MPN<sup>+</sup> partner. The complex regulation mechanisms of the DUB activity among which hetero-dimerisation warrant further studies to understand the isopeptidase activity control of these large assemblies. Furthermore, members of the family such as MPND and MYSM1 are yet to be studied in details.

## CONFLICT OF INTEREST

The author(s) confirm that this article content has no conflicts of interest.

## ACKNOWLEDGEMENTS

M. Birol was awarded a PhD fellowship at the University Montpellier II by the Doctoral school CBS2. A. Echaliér was supported by a chaire mixte University Montpellier I/CNRS. Work in the author's lab is supported by research grants

from the CNRS PEPS 2009 'Biochemical and structural characterisation of multi-protein complexes implicated in the regulation of the ubiquitin-proteasome pathway', GEFLUC Languedoc-Roussillon 2012 award and a BUSR Université Montpellier I grant. A. Echaliér's laboratory is an associated member in the Marie Curie UPStream ITN.

## ABBREVIATIONS

AMSH	= Associated molecule with the SH3 domain of STAM
AMSH-LP	= AMSH-like protein
BRCC36	= BRCA1/BRCA2-containing complex subunit 36
BRISC	= BRCC36 isopeptidase complex
COP9	= Constitutive photomorphogenesis 9
CRLs	= Cullin RING E3 ligases
DDR	= DNA damage response
DUB	= Deubiquitinylase
eIF3	= Eukaryotic translation initiation factor 3
eIF3f	= Eukaryotic translation initiation factor 3 subunit f
eIF3h	= Eukaryotic translation initiation factor 3 subunit h
IFNAR1	= Interferon receptor chain 1
JAMM	= Jab1-MPN-Mov34
MPN	= Mpr1/Pad1 N-terminal
MYSM1	= Myb-like, SWIRM, and MPN domain-containing protein 1
Nedd8	= Neural precursor cell expressed developmentally down-regulated protein 8
p/CAF	= Histone acetyltransferase (HAT) p300/CBP-associated factor
PDB	= Protein Data Bank
RP	= Regulatory particle
SANT	= Swi3, Ada2, N-Cor, and TFIIB
SHMT	= Serine hydroxymethyltransferase
snRNP	= Small nuclear ribonucleoprotein particle
STAM	= Signal Transducing Adaptor Molecule
SWIRM	= Swi3, Rsc8, and Moira
UPS	= Ubiquitin-proteasome system

## REFERENCES

- [1] Nijman, S.M.; Luna-Vargas, M.P.; Velds, A.; Brummelkamp, T.R.; Dirac, A.M.; Sixma, T.K.; Bernards, R. A genomic and functional inventory of deubiquitinating enzymes. *Cell*, **2005**, *123*, 773-786.
- [2] Ambroggio, X.I.; Rees, D.C.; Deshaies, R.J. JAMM: a metalloprotease-like zinc site in the proteasome and signalosome. *PLoS Biol.*, **2004**, *2*, E2.
- [3] Tran, H.J.; Allen, M.D.; Lowe, J.; Bycroft, M. Structure of the Jab1/MPN domain and its implications for proteasome function. *Biochemistry*, **2003**, *42*, 11460-11465.

- [4] Pena, V.; Liu, S.; Bujnicki, J.M.; Luhrmann, R.; Wahl, M.C. Structure of a multipartite protein-protein interaction domain in splicing factor prp8 and its link to retinitis pigmentosa. *Mol. Cell.*, **2007**, *25*, 615-624.
- [5] Sanches, M.; Alves, B.S.; Zanchin, N.I.; Guimaraes, B.G. The crystal structure of the human Mov34 MPN domain reveals a metal-free dimer. *J. Mol. Biol.*, **2007**, *370*, 846-855.
- [6] Zhang, L.; Shen, J.; Guarnieri, M.T.; Heroux, A.; Yang, K.; Zhao, R. Crystal structure of the C-terminal domain of splicing factor Prp8 carrying retinitis pigmentosa mutants. *Protein Sci.*, **2007**, *16*, 1024-1031.
- [7] Sato, Y.; Yoshikawa, A.; Yamagata, A.; Mimura, H.; Yamashita, M.; Ookata, K.; Nureki, O.; Iwai, K.; Komada, M.; Fukai, S. Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains. *Nature*, **2008**, *455*, 358-362.
- [8] Davies, C.W.; Paul, L.N.; Kim, M.I.; Das, C. Structural and thermodynamic comparison of the catalytic domain of AMSH and AMSH-LP: nearly identical fold but different stability. *J. Mol. Biol.*, **2011**, *413*, 416-429.
- [9] Zhang, H.; Gao, Z.Q.; Wang, W.J.; Liu, G.F.; Shtykova, E.V.; Xu, J.H.; Li, L.F.; Su, X.D.; Dong, Y.H. The crystal structure of the MPN domain from the COP9 signalosome subunit CSN6. *FEBS Lett.*, **2012**, *586*, 1147-1153.
- [10] Echallier, A.; Pan, Y.; Birol, M.; Tavernier, N.; Pintard, L.; Hoh, F.; Ebel, C.; Galoppe, N.; Claret, F.X.; Dumas, C. Insights into the regulation of the human COP9 signalosome catalytic subunit, CSN5/Jab1. *Proc. Natl. Acad. Sci. U S A*, **2013**, *110*, 1273-1278.
- [11] Galej, W.P.; Oubridge, C.; Newman, A.J.; Nagai, K. Crystal structure of Prp8 reveals active site cavity of the spliceosome. *Nature*, **2013**, *493*, 638-643.
- [12] Mozaffari-Jovin, S.; Wandersleben, T.; Santos, K.F.; Will, C.L.; Luhrmann, R.; Wahl, M.C. Inhibition of RNA Helicase Br2 by the C-Terminal Tail of the Spliceosomal Protein Prp8. *Science*, **2013**, *341*(6141), 80-84.
- [13] Nguyen, T.H.; Li, J.; Galej, W.P.; Oshikane, H.; Newman, A.J.; Nagai, K. Structural Basis of Br2-Prp8 Interactions and Implications for U5 snRNP Biogenesis and the Spliceosome Active Site. *Structure*, **2013**, *21*(6), 910-919.
- [14] Weber, G.; Cristao, V.F.; Santos, K.F.; Jovin, S.M.; Heroven, A.C.; Holton, N.; Luhrmann, R.; Beggs, J.D.; Wahl, M.C. Structural basis for dual roles of Aar2p in U5 snRNP assembly. *Genes Dev.*, **2013**, *27*, 525-540.
- [15] Grainger, R.J.; Beggs, J.D. Prp8 protein: at the heart of the spliceosome. *RNA*, **2005**, *11*, 533-557.
- [16] Zhu, P.; Zhou, W.; Wang, J.; Puc, J.; Ohgi, K.A.; Erdjument-Bromage, H.; Tempst, P.; Glass, C.K.; Rosenfeld, M.G. A histone H2A deubiquitinase complex coordinating histone acetylation and H1 dissociation in transcriptional regulation. *Mol. Cell*, **2007**, *27*, 609-621.
- [17] Cooper, E.M.; Cutcliffe, C.; Kristiansen, T.Z.; Pandey, A.; Pickart, C.M.; Cohen, R.E. K63-specific deubiquitination by two JAMM/MPN+ complexes: BRISC-associated Brcc36 and proteasomal Pohl1. *EMBO J.*, **2009**, *28*, 621-631.
- [18] Patterson-Fortin, J.; Shao, G.; Bretscher, H.; Messick, T.E.; Greenberg, R.A. Differential regulation of JAMM domain deubiquitinating enzyme activity within the RAP80 complex. *J. Biol. Chem.*, **2010**, *285*, 30971-30981.
- [19] Sharon, M.; Mao, H.; Boeri Erba, E.; Stephens, E.; Zheng, N.; Robinson, C.V. Symmetrical modularity of the COP9 signalosome complex suggests its multifunctionality. *Structure*, **2009**, *17*, 31-40.
- [20] Sharon, M.; Tavernier, T.; Ambroggio, X.I.; Deshaies, R.J.; Robinson, C.V. Structural organization of the 19S proteasome lid: insights from MS of intact complexes. *PLoS Biol.*, **2006**, *4*, e267.
- [21] Zhou, M.; Sandercock, A.M.; Fraser, C.S.; Ridlova, G.; Stephens, E.; Schenauer, M.R.; Yokoi-Fong, T.; Barsky, D.; Leary, J.A.; Hershey, J.W.; Doudna, J.A.; Robinson, C.V. Mass spectrometry reveals modularity and a complete subunit interaction map of the eukaryotic translation factor eIF3. *Proc. Natl. Acad. Sci. U S A*, **2008**, *105*, 18139-18144.
- [22] Pick, E.; Pintard, L. In the land of the rising sun with the COP9 signalosome and related Zomes. Symposium on the COP9 signalosome, Proteasome and eIF3. *EMBO Rep.*, **2009**, *10*, 343-348.
- [23] Bellare, P.; Kutach, A.K.; Rines, A.K.; Guthrie, C.; Sontheimer, E.J. Ubiquitin binding by a variant Jab1/MPN domain in the essential pre-mRNA splicing factor Prp8p. *RNA*, **2006**, *12*, 292-302.
- [24] Cope, G.A.; Suh, G.S.; Aravind, L.; Schwarz, S.E.; Zipursky, S.L.; Koonin, E.V.; Deshaies R.J. Role of predicted metalloprotease motif of Jab1/Csn5 in cleavage of Nedd8 from Cul1. *Science*, **2002**, *298*, 608-611.
- [25] Hepowit, N.L.; Uthandi, S.; Miranda, H.V.; Toniutti, M.; Prunetti, L.; Olivarez, O.; De Vera, I.M.; Fanucci, G.E.; Chen, S.; Maupin-Furlow, J.A. Archaeal JAB1/MPN/MOV34 metalloenzyme (HvJAMM1) cleaves ubiquitin-like small archaeal modifier proteins (SAMPs) from protein-conjugates. *Mol. Microbiol.*, **2012**, *86*(4), 971-87.
- [26] McCullough, J.; Clague, M.J.; Urbe, S. AMSH is an endosome-associated ubiquitin isopeptidase. *J. Cell Biol.*, **2004**, *166*, 487-492.
- [27] Moretti, J.; Chastagner, P.; Gastaldello, S.; Heuss, S.F.; Dirac, A.M.; Bernards, R.; Masucci, M.G.; Israël, A.; Brou, C. The translation initiation factor 3f (eIF3f) exhibits a deubiquitinase activity regulating Notch activation. *PLoS Biol.*, **2010**, *8*, e1000545.
- [28] Nakamura, M.; Tanaka, N.; Kitamura, N.; Komada, M. Clathrin anchors deubiquitinating enzymes, AMSH and AMSH-like protein, on early endosomes. *Genes Cells*, **2006**, *11*, 593-606.
- [29] Verma, R.; Aravind, L.; Oania, R.; McDonald, W.H.; Yates, J.R. 3rd.; Koonin, E.V.; Deshaies, R.J. Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. *Science*, **2002**, *298*, 611-615.
- [30] Komander, D. The emerging complexity of protein ubiquitination. *Biochem. Soc. Trans.*, **2009**, *37*, 937-953.
- [31] Marchionne, R.; Leibovitch, S.A.; Lenormand, J.L. The translational factor eIF3f: the ambivalent eIF3 subunit. *Cell Mol. Life Sci.*, **2013**, *70*, 3603-3616.
- [32] Will, C.L.; Luhrmann, R. Spliceosome structure and function. *Cold Spring Harb. Perspect. Biol.*, **2011**, *3*(7), pii: a003707.
- [33] Shackleford, T.J.; Claret, F.X. JAB1/CSN5: a new player in cell cycle control and cancer. *Cell Div.*, **2010**, *5*, 26.
- [34] Zhang, S.N.; Pei, D.S.; Zheng, J.N. The COP9 signalosome subunit 6 (CSN6): a potential oncogene. *Cell Div.*, **2013**, *8*, 14.
- [35] Nandakumar, V.; Chou, Y.; Zang, L.; Huang, X.F.; Chen, S.Y. Epigenetic control of natural killer cell maturation by histone H2A deubiquitinase, MYSM1. *Proc. Natl. Acad. Sci. U S A*, **2013**, *110*, E3927-3936.
- [36] Maytal-Kivity, V.; Pick, E.; Piran, R.; Hofmann, K.; Glickman, M.H. The COP9 signalosome-like complex in *S. cerevisiae* and links to other PCI complexes. *Int. J. Biochem. Cell Biol.*, **2003**, *35*, 706-715.
- [37] Pearce, M.J.; Mintseris, J.; Ferreyra, J.; Gygi, S.P.; Darwin, K.H. Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. *Science*, **2008**, *322*, 1104-1107.
- [38] Maytal-Kivity, V.; Reis, N.; Hofmann, K.; Glickman, M.H. MPN+, a putative catalytic motif found in a subset of MPN domain proteins from eukaryotes and prokaryotes, is critical for Rpn11 function. *BMC Biochem.*, **2002**, *3*, 28.
- [39] Lyapina, S.; Cope, G.; Shevchenko, A.; Serino, G.; Tsuge, T.; Zhou, C.; Wolf, D.A.; Wei, N.; Shevchenko, A.; Deshaies, R.J. Promotion of NEDD-CUL1 conjugate cleavage by COP9 signalosome. *Science*, **2001**, *292*, 1382-1385.
- [40] Braus, G.H.; Irniger, S.; Bayram, O. Fungal development and the COP9 signalosome. *Curr. Opin. Microbiol.*, **2010**, *13*, 672-676.
- [41] Pick, E.; Golan, A.; Zimble, J.Z.; Guo, L.; Sharaby, Y.; Tsuge, T.; Hofmann, K.; Wei, N. The minimal deneddylase core of the COP9 signalosome excludes the Csn6 MPN- domain. *PLoS One*, **2012**, *7*, e43980.
- [42] Fukumoto, A.; Tomoda, K.; Kubota, M.; Kato, J.Y.; Yoneda-Kato, N. Small Jab1-containing subcomplex is regulated in an anchorage- and cell cycle-dependent manner, which is abrogated by ras transformation. *FEBS Lett.*, **2005**, *579*, 1047-1054.
- [43] Tomko, R.J. Jr.; Hochstrasser, M. Molecular architecture and assembly of the eukaryotic proteasome. *Annu. Rev. Biochem.*, **2013**, *82*, 415-445.
- [44] Matyskiela, M.E.; Martin, A. Design principles of a universal protein degradation machine. *J. Mol. Biol.*, **2013**, *425*, 199-213.
- [45] Singh, R.K.; Zerath, S.; Kleifeld, O.; Scheffner, M.; Glickman, M.H.; Fushman, D. Recognition and cleavage of related to ubiquitin 1 (Rub1) and Rub1-ubiquitin chains by components of the ubiquitin-proteasome system. *Mol. Cell Proteomics*, **2012**, *11*, 1595-1611.

- [46] Rinaldi, T.; Pick, E.; Gambadoro, A.; Zilli, S.; Maytal-Kivity, V.; Frontali, L.; Glickman, M.H. Participation of the proteasomal lid subunit Rpn11 in mitochondrial morphology and function is mapped to a distinct C-terminal domain. *Biochem. J.*, **2004**, *381*, 275-285.
- [47] Beck, F.; Unverdorben, P.; Bohn, S.; Schweitzer, A.; Pfeifer, G.; Sakata, E.; Nickell, S.; Plitzko, J.M.; Villa, E.; Baumeister, W.; Förster, F. Near-atomic resolution structural model of the yeast 26S proteasome. *Proc. Natl. Acad. Sci. U S A*, **2012**, *109*, 14870-14875.
- [48] da Fonseca, P.C.; He, J.; Morris, E.P. Molecular model of the human 26S proteasome. *Mol. Cell*, **2012**, *46*, 54-66.
- [49] Estrin, E.; Lopez-Blanco, J.R.; Chacon, P.; Martin, A. Formation of an intricate helical bundle dictates the assembly of the 26S proteasome lid. *Structure*, **2013**, *21*, 1624-1635.
- [50] Lasker, H.; Forster, F.; Bohn, S.; Walzthoeni, T.; Villa, E.; Unverdorben, P.; Beck, F.; Aebersold, R.; Sali, A.; Baumeister, W. Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proc. Natl. Acad. Sci. U S A*, **2012**, *109*, 1380-1387.
- [51] Zheng, H.; Gupta, V.; Patterson-Fortin, J.; Bhattacharya, S.; Katlinski, K.; Wu, J.; Varghese, B.; Carbone, C.J.; Aressy, B.; Fuchs, S.Y.; Greenberg, R.A. A BRISC-SHMT complex deubiquitinates IFNAR1 and regulates interferon responses. *Cell Rep.*, **2013**, *5*, 180-193.
- [52] Wang, B.; Matsuoka, S.; Ballif, B.A.; Zhang, D.; Smogorzewska, A.; Gygi, S.P.; Elledge, S.J. Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science*, **2007**, *316*, 1194-1198.
- [53] Kikuchi, K.; Ishii, N.; Asao, H.; Sugamura, K. Identification of AMSH-LP containing a Jab1/MPN domain metalloenzyme motif. *Biochem. Biophys. Res. Commun.*, **2003**, *306*, 637-643.
- [54] Jiang, X.X.; Nguyen, Q.; Chou, Y.; Wang, T.; Nandakumar, V.; Yates, P.; Jones, L.; Wang, L.; Won, H.; Lee, H.R.; Jung, J.U.; Müschen, M.; Huang, X.F.; Chen, S.Y. Control of B cell development by the histone H2A deubiquitinase MYSM1. *Immunity*, **2011**, *35*, 883-896.
- [55] Wang, T.; Nandakumar, V.; Jiang, X.X.; Jones, L.; Yang, A.G.; Huang, X.F.; Chen, S.Y. The control of hematopoietic stem cell maintenance, self-renewal, and differentiation by Mysm1-mediated epigenetic regulation. *Blood*, **2013**, *122*, 2812-2822.
- [56] Sun, C.; Todorovic, A.; Querol-Audi, J.; Bai, Y.; Villa, N.; Snyder, M.; Ashchyan, J.; Lewis, C.S.; Hartland, A.; Gradia, S.; Fraser, C.S.; Doudna, J.A.; Nogales, E.; Cate, J.H. Functional reconstitution of human eukaryotic translation initiation factor 3 (eIF3). *Proc. Natl. Acad. Sci. U S A*, **2011**, *108*, 20473-20478.
- [57] Weber, G.; Cristao, V.F.; de, L.A.F.; Santos, K.F.; Holton, N.; Rappalber, J.; Beggs, J.D.; Wahl, M.C. Mechanism for Aar2p function as a U5 snRNP assembly factor. *Genes Dev.*, **2011**, *25*, 1601-1612.
- [58] Enchev, R.I.; Schreiber, A.; Beuron, F.; Morris, E.P. Structural insights into the COP9 signalosome and its common architecture with the 26S proteasome lid and eIF3. *Structure*, **2010**, *18*, 518-527.
- [59] Enchev, R.I.; Scott, D.C.; da Fonseca, P.C.; Schreiber, A.; Monda, J.K.; Schulman, B.A.; Peter, M.; Morris, E.P. Structural basis for a reciprocal regulation between SCF and CSN. *Cell Rep.*, **2012**, *2*, 616-627.
- [60] Querol-Audi, J.; Sun, C.; Vogan, J.M.; Smith, M.D.; Gu, Y.; Cate, J.H.; Nogales, E. Architecture of human translation initiation factor 3. *Structure*, **2013**, *21*, 920-928.
- [61] Komander, D. Mechanism, specificity and structure of the deubiquitinases. *Subcell. Biochem.*, **2010**, *54*, 69-87.
- [62] Komander, D.; Clague, M.J.; Urbe, S. Breaking the chains: structure and function of the deubiquitinases. *Nat. Rev. Mol. Cell Biol.*, **2009**, *10*, 550-563.
- [63] Emberley, E.D.; Mosadeghi, R.; Deshaies, R.J. Deconjugation of Nedd8 from Cull1 is directly regulated by Skp1-F-box and substrate, and the COP9 signalosome inhibits deneddylated SCF by a noncatalytic mechanism. *J. Biol. Chem.*, **2012**, *287*, 29679-29689.
- [64] Ashkenazy, H.; Erez, E.; Martz, E.; Pupko, T.; Ben-Tal, N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.*, **2010**, *38*, W529-533.
- [65] Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.*, **2007**, *372*, 774-797.
- [66] Fu, H.; Reis, N.; Lee, Y.; Glickman, M.H.; Vierstra, R.D. Subunit interaction maps for the regulatory particle of the 26S proteasome and the COP9 signalosome. *EMBO J.*, **2001**, *20*, 7096-7107.
- [67] Peng, Z.; Serino, G.; Deng, X.W. Molecular characterization of subunit 6 of the COP9 signalosome and its role in multifaceted developmental processes in Arabidopsis. *Plant Cell*, **2001**, *13*, 2393-2407.
- [68] Sowa, M.E.; Bennett, E.J.; Gygi, S.P.; Harper, J.W. Defining the human deubiquitinating enzyme interaction landscape. *Cell*, **2009**, *138*, 389-403.
- [69] Wang, B.; Elledge, S.J. Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brcal/Brc36 complex in response to DNA damage. *Proc. Natl. Acad. Sci. U S A*, **2007**, *104*, 20759-20763.
- [70] Cooper, E.M.; Boeke, J.D.; Cohen, R.E. Specificity of the BRISC deubiquitinating enzyme is not due to selective binding to Lys63-linked polyubiquitin. *J. Biol. Chem.*, **2010**, *285*, 10344-10352.

## **Paper 3**

**Structural and Biochemical Characterization of the COP9 Signalosome CSN5/CSN6 Heterodimer**



# Structural and Biochemical Characterization of the Cop9 Signalosome CSN5/CSN6 Heterodimer

Melissa Birol<sup>1,2</sup>, Radoslav Ivanov Enchev<sup>3</sup>, André Padilla<sup>1,2</sup>, Florian Stengel<sup>4</sup>, Ruedi Aebersold<sup>4,5</sup>, Stéphane Betzi<sup>6</sup>, Yinshan Yang<sup>1,2</sup>, François Hoh<sup>1,2</sup>, Matthias Peter<sup>3</sup>, Christian Dumas<sup>1,2</sup>, Aude Echallier<sup>1,2\*</sup>

**1** Centre de Biochimie Structurale, Unité Mixte de Recherche (UMR) 5048, Centre National de Recherche Scientifique (CNRS), Université Montpellier 1 (UM1), Université Montpellier 2 (UM2), Montpellier, France, **2** Institut national de la santé et de la recherche médicale (INSERM) U1054, Paris, France, **3** ETH Zurich, Department of Biology, Institute of Biochemistry, Zurich, Switzerland, **4** ETH Zurich, Department of Biology, Institute of Molecular Systems Biology, Zurich, Switzerland, **5** Faculty of Science, University of Zurich, Zurich, Switzerland, **6** Centre de Recherche en Cancérologie de Marseille, Centre de Biochimie Structurale, Unité Mixte de Recherche (UMR) 7258, Institut national de la santé et de la recherche médicale (INSERM) U1068, Institut Paoli-Calmettes, Aix-Marseille Université UM105, Marseille, France

## Abstract

The Cop9 signalosome complex (CSN) regulates the functional cycle of the major E3 ubiquitin ligase family, the cullin RING E3 ubiquitin ligases (CRLs). Activated CRLs are covalently modified by the ubiquitin-like protein Nedd8 (neural precursor cell expressed developmentally down-regulated protein 8). CSN serves an essential role in myriad cellular processes by reversing this modification through the isopeptidase activity of its CSN5 subunit. CSN5 alone is inactive due to an auto-inhibited conformation of its catalytic domain. Here we report the molecular basis of CSN5 catalytic domain activation and unravel a molecular hierarchy in CSN deneddylation activity. The association of CSN5 and CSN6 MPN (for Mpr1/Pad1 N-terminal) domains activates its isopeptidase activity. The CSN5/CSN6 module, however, is inefficient in CRL deneddylation, indicating a requirement of further elements in this reaction such as other CSN subunits. A hybrid molecular model of CSN5/CSN6 provides a structural framework to explain these functional observations. Docking this model into a published CSN electron density map and using distance constraints obtained from cross-linking coupled to mass-spectrometry, we find that the C-termini of the CSN subunits could form a helical bundle in the centre of the structure. They likely play a key scaffolding role in the spatial organization of CSN and precise positioning of the dimeric MPN catalytic core.

**Citation:** Birol M, Enchev RI, Padilla A, Stengel F, Aebersold R, et al. (2014) Structural and Biochemical Characterization of the Cop9 Signalosome CSN5/CSN6 Heterodimer. PLoS ONE 9(8): e105688. doi:10.1371/journal.pone.0105688

**Editor:** Elena Papaleo, University of Copenhagen, Denmark

**Received:** June 12, 2014; **Accepted:** July 23, 2014; **Published:** August 21, 2014

**Copyright:** © 2014 Birol et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability:** The authors confirm that all data underlying the findings are fully available without restriction. The files related to the crystal structure described in this paper are available from the PDB database (accession number 4QFT). All remaining relevant data are within the paper and its Supporting Information files.

**Funding:** Funding provided by Swiss National Science Foundation: European Research Council senior award: MP, ETH-Zurich: MP RA, Marie Curie: RIE, Sir Henry Wellcome Fellow: FS (Wellcome Trust Grant 095951), European Research Council advanced grant Proteomics v3.0 (233226): RA, CBS2/UM2 PhD studentship: MB, University Montpellier I: AE, Centre national de la recherche scientifique: AE and GEFLUC Languedoc Roussillon: AE. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* Email: aude.echallier-glazer@cbs.cnrs.fr

‡ Current address: Departments of Biochemistry and of Cancer Studies and Molecular Medicine, University of Leicester, Leicester, United Kingdom

## Introduction

The ubiquitin-proteasome system is implicated in virtually all functions of eukaryotic living cells. The covalent attachment of ubiquitin molecules, ubiquitylation, requires the concerted intervention of the distinct E1, E2 and E3 enzymes. The most prominent E3 ubiquitin ligase family comprises the cullin RING E3 ubiquitin ligases (CRLs). CRLs are important regulators of cellular homeostasis, division, and responses to various cellular insults [1]. Their significance is further highlighted by the deregulation of some CRL elements in human diseases and, in particular, in cancers [2,3].

CRLs are built around a cullin scaffold, which associates with a substrate-specific adaptor and a RING finger protein [4]. CRL activity is tightly regulated to ensure the timely and specific substrate ubiquitylation [1]. The covalent attachment of the ubiquitin-like (UBL) molecule Nedd8 (neural precursor cell

expressed developmentally down-regulated protein 8) to cullins, termed neddylation, activates their ubiquitylation activity by stabilising an activated conformation of the RING subunit [5,6].

The Cop9 signalosome (CSN) is an eight-subunit complex that deneddylates cullins [7]. Six of its subunits (CSN1 to CSN4, CSN7, CSN8) contain a PCI (for Proteasome Cop9 eIF3) domain and two (CSN5/6) contain an MPN (for Mpr1/Pad1 N-terminal) domain. The PCI subunits serve a scaffolding function and, especially through CSN2, help recruit the neddylated CRL substrates [8]. Structurally, the PCI-containing subunits arrange in a horseshoe-like shape, juxtaposed with the two MPN subunits. This architecture is strikingly similar to that of the 26S proteasome lid and the eukaryotic translation initiation factor eIF3 [8,9,10,11,12,13].

Among the CSN MPN-containing subunits, CSN6 lacks a functional catalytic site and belongs to the MPN<sup>−</sup> class [14,15,16].



In contrast, CSN5, also known as Jun-activatory binding protein 1 (Jab1) carries the catalytically competent MPN<sup>+</sup>/JAMM (Jab1-MPN-Mov34) motif [15,17]. As defined in the AMSH (Associated Molecule with the SH3 domain of STAM)-like protein (AMSH-LP) [18], the MPN domain of CSN5 contains two insertions, namely Insertion-1 (Ins-1; residues 97–131 in human CSN5) and Insertion-2 (Ins-2; residues 197–219 in human CSN5). These regions usually contribute to the regulation of the isopeptidase activity. For AMSH-LP, the Ins-1 region plays a role in the binding and competent positioning of the distal ubiquitin [18]. CSN5 by itself is found in an inactive state and the transition from this stand-alone auto-inhibited to the active form probably requires conformational changes of the Ins-1 region [17,19]. This is reminiscent of Rpn11 and BRCC36 (for BRCA1 (for Breast Cancer 1)/BRCA2 (for Breast Cancer 2)-Containing Complex subunit 36) found in the 26S proteasome lid and BRCC36-containing complexes, respectively, that are also inactive in the stand-alone form [20]. Recently, the crystal structure of Rpn11/Rpn8 MPN domains brought important insights into the dimerisation of an MPN<sup>+</sup>/JAMM enzyme with an MPN<sup>−</sup> subunit [21,22], pointing towards the mechanism underlying the catalytic activation of inactive MPN<sup>+</sup>/JAMM enzymes upon integration in higher order assemblies.

Although the structural and functional understanding of the CSN/CRL interplay has significantly grown over the recent years, important questions concerning the association of the MPN domains and the regulation of CSN activity remain. Here we report that CSN6 N-terminal domain containing its MPN domain can form a stable heterodimer with CSN5 catalytic domain *in vitro*. Biophysical and activity measurements indicate that CSN6 MPN domain stimulates CSN5's isopeptidase activity. Comparison with CSN on synthetic substrates reveals that both CSN and the dimeric MPN module display robust isopeptidase activity towards C-terminal Nedd8 derivatives, but that CSN is a more efficient deneddylase towards Cullin1-Nedd8 than the CSN5/CSN6 MPN complex. This suggests that the MPN<sup>−</sup> subunit would contribute significantly to the catalytic activity of the human CSN, but efficient substrate recruitment would require additional elements such as other CSN subunits. Moreover, we provide a structural context for these observations, combining X-ray crystallography, NMR and cross-linking coupled to MS (CX-MS) derived data and flexible molecular docking to produce a molecular model of the CSN5/CSN6 MPN heterodimer and re-interpret the molecular model of CSN (8). This analysis suggests that a major part of the subunit-subunit interactions in CSN are mediated by a helical bundle, composed of the C-termini of CSN's subunits, further highlighting its structural similarity to the 26S proteasome lid [23]. Our work indicates that, in addition to the importance of their MPN domains for catalysis, the C-terminal regions of CSN5 and of CSN6 contribute to their anchoring and precise positioning in the CSN assembly.

## Experimental Procedures

### Protein production

The expression and purification of human CSN5 1–257 fragment, in the WT and mutant forms were performed as previously described in [19]. The human CSN6 cDNA was subcloned in pGEX-6P1 vector (Novagen). The CSN6 31–211 fragment was found to be solubly expressed in bacteria and was selected for further work. The 1–257 and 31–211 fragments of CSN5 and CSN6, corresponding to their MPN domains together with N- and C-terminal appendices for CSN5 and a C-terminal extension for CSN6, as illustrated in Figure 1A, are referred to as

CSN5<sup>ΔC</sup> and CSN6<sup>ΔC</sup>, respectively, in the rest of the manuscript. Expression of CSN6<sup>ΔC</sup> and its variants was performed in *E. coli* BL21pLysS cells (Novagen) in LB or <sup>15</sup>NH<sub>4</sub>Cl-supplemented M9 minimal expression medium, as required for NMR experiments. Protein expression was carried out overnight at 18°C and the following purification buffer (20 mM Tris pH 7.5, 150 mM NaCl, 0.01% monothioglycerol) was used throughout the purification. A standard purification protocol for GST-tagged proteins was used, including a first glutathione sepharose affinity step, a size exclusion chromatography step and a second affinity step to remove contaminating GST from the purified protein. Purified CSN6<sup>ΔC</sup> was concentrated to 8 mg mL<sup>−1</sup> and stored at −80°C until further use. Production of pro-Nedd8 from the pOPIN-E-pro-nedd8 vector (from M. Banfield), was carried out following the protocol described in [24]. Site directed mutagenesis performed in this work were done with the QuikChange Lightning Site-Directed mutagenesis kit (Stratagene) and verified by DNA sequencing. Production and purification procedures for the CSN and Cullin1-Nedd8/Rbx1 were carried out following the protocols described in [8].

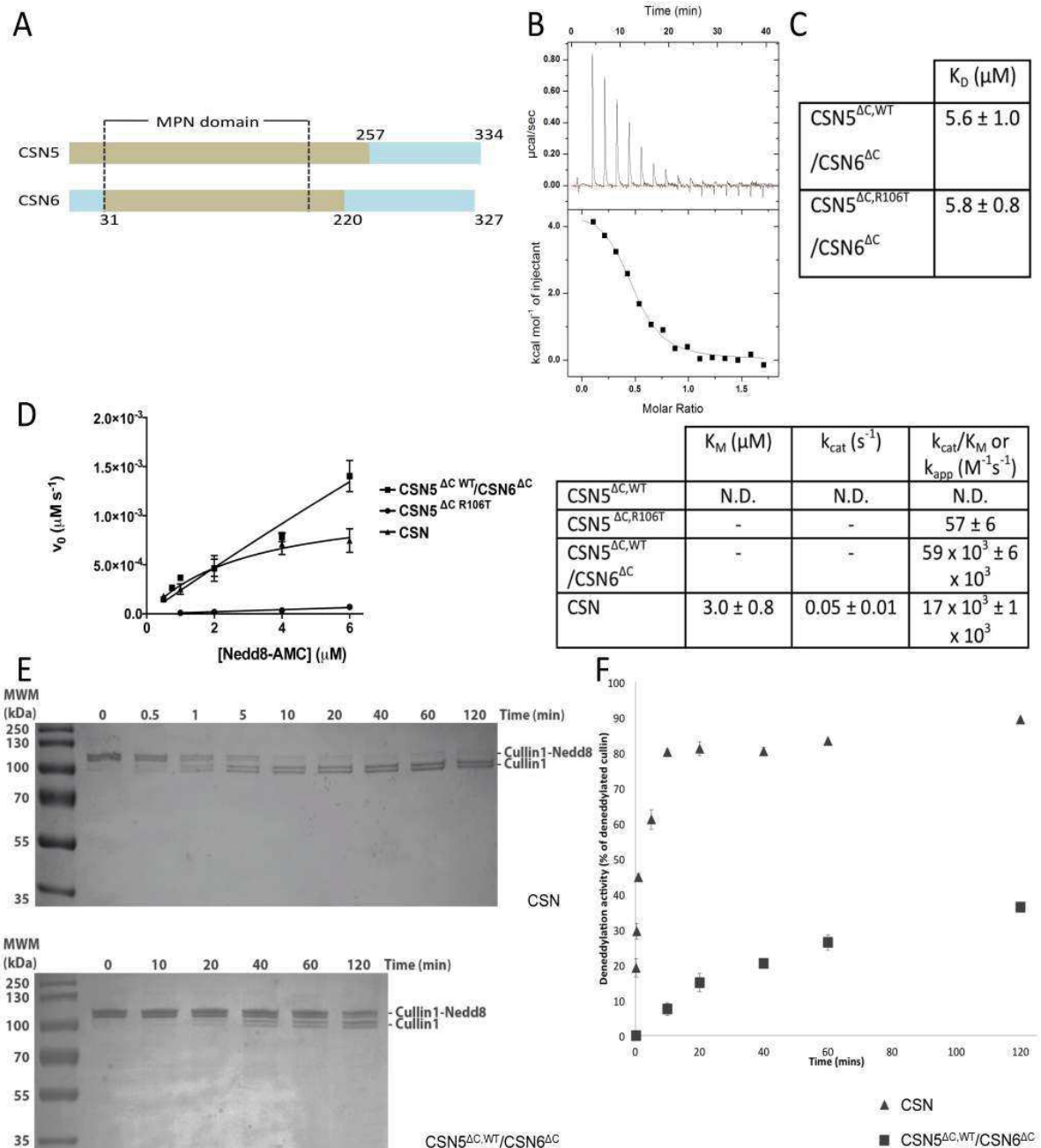
### Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) experiments were carried out on a MicroCal ITC<sub>200</sub> microcalorimeter (GE Healthcare, Piscataway, NJ) at 20°C. The protein samples were all buffer exchanged using PD-10 desalting column (GE Healthcare) into the ITC buffer (20 mM Na MES pH 6.5, 75 mM NaCl). Protein concentrations were measured using a NanoDrop 1000 spectrophotometer (Thermo Scientific, Wilmington, DE), with the following extinction coefficients: 56,840 M<sup>−1</sup>cm<sup>−1</sup> and 18,450 M<sup>−1</sup>cm<sup>−1</sup> for CSN5<sup>ΔC</sup> and CSN6<sup>ΔC</sup>, respectively. Titration of CSN5<sup>ΔC</sup> (WT and variants; 100 μM) in the cell (200 μL) was performed by sequential addition of CSN6<sup>ΔC</sup> (WT and variants; 1 mM; 30 injections of 1.8 μL). Integrated raw ITC data were fitted to a one site nonlinear least squares fit model using the MicroCal Origin plugin as implemented in Origin 9.1 (Origin Lab) after the control experiments (titration of the ligand from the syringe into the buffer) were subtracted.

### Fluorescence anisotropy

To obtain quantitative information about the affinity of CSN5<sup>ΔC</sup> to Nedd8, fluorescence anisotropy measurements were performed at 28°C, using a TECAN safire<sup>2</sup> plate reader (Tecan). The experiments were performed in 20 mM Hepes pH 7.5 and 75 mM NaCl. Alexa Fluor 488 fluorophore (Life Technologies) was used to label Nedd8 through a standard amine coupling procedure. The fluorescence was measured using excitation and emission wavelengths of 495 nm and 519 nm, respectively. CSN5<sup>ΔC,WT or R106T</sup> at varying concentrations (0–600 μM) was incubated in the presence of 4 nM Alexa Fluor 488-Nedd8. The same procedure was followed in the presence of CSN6<sup>ΔC</sup> where a 1:1 ratio of CSN5<sup>ΔC</sup>/CSN6<sup>ΔC</sup> was used to make the complex at the indicated concentrations. For each data point the measurement was done in triplicates and the recorded data corresponds to the average anisotropy. The data obtained were plotted and analysed using the IgorPro software. The Hill Equation was used for fitting the data in order to be able to deduct an estimate of the apparent equilibrium dissociation constant (K<sub>D</sub>) of each complex between the protein or protein complex and Alexa Fluor 488-Nedd8. As there exists no evidence of cooperativity, in all cases, the data were fit with values of cooperativity equal to 1.





**Figure 1. CSN6<sup>ΔC</sup> activates CSN5<sup>ΔC</sup>.** (A) Studied protein fragments (brown). The MPN domain boundaries are indicated with dotted lines. (B) Binding of CSN5<sup>ΔC</sup> and CSN6<sup>ΔC</sup> measured by ITC. Titration experiments were carried out by injecting concentrated CSN6<sup>ΔC</sup> solution into 100  $\mu\text{M}$  CSN5 variants. (C)  $K_D$  values of CSN5<sup>ΔC</sup>/CSN6<sup>ΔC</sup> pairs obtained from ITC data. (D) Activity towards Nedd8-AMC hydrolysis rate plot versus substrate concentration (200 nM CSN5<sup>ΔC</sup> alone; 4 nM CSN5<sup>ΔC</sup>/CSN6<sup>ΔC</sup> complex; 23 nM CSN), with data fitted using the Michaelis-Menten equation. Below: Associated kinetic parameters of the Nedd8-AMC cleavage derived from the left plot. In the CSN5<sup>ΔC,R106T</sup> and CSN5<sup>ΔC,WT</sup>/CSN6<sup>ΔC</sup> experiments, substrate saturation conditions were not reached and  $k_{app}$  values were extracted from linear fitting of the  $k_{obs}$  versus substrate concentration data. (E) Time-course deneddylation of Cullin1-Nedd8/Rbx1 (1  $\mu\text{M}$ ) by the CSN (4 nM; Top) or CSN5<sup>ΔC,WT</sup>/CSN6<sup>ΔC</sup> (4 nM; Bottom). Samples from each time point were analysed on Coomassie stained 15% Tris-tricine SDS-PAGE. Quantification of the cullin deneddylation was carried out. The bands corresponding to Cullin1-Nedd8 and to Cullin1 show doublet that arise from Cullin1 degradation. (F) Deneddylation activity on the Cullin1-Nedd8/Rbx1. Substrate hydrolysis (1  $\mu\text{M}$ ) is shown as a function of time for 4 nM enzyme (CSN, triangle; CSN5<sup>ΔC,WT</sup>/CSN6<sup>ΔC</sup>, square). For context, CSN5<sup>ΔC,R106T</sup> alone at 900 nM concentration produces less than 30% of deneddylated Cullin1 for 120 min reaction (data point not shown for clarity). Data information: N.D., not significant; -, parameters not measured; error bars = s.d.; experiments done at least twice.

doi:10.1371/journal.pone.0105688.g001

### Isopeptidase assays

Activity assays have been carried out on synthetic and physiological substrates. A synthetic substrate, namely Nedd8-AMC (Nedd8 protein conjugated to a 7-amino-4-methyl-coumarin molecule (AMC); from Boston Biochem.), and one physiological substrate (Cullin1-Nedd8/Rbx1) were used to assess the catalytic activities of different proteins and complexes used in this study. For the Nedd8-AMC substrate, initial rate measurements were carried out at 37°C by following the increase of fluorescence intensity ( $\lambda_{\text{excitation}} = 380 \text{ nm}$ ;  $\lambda_{\text{emission}} = 460 \text{ nm}$ ), in non-binding surface, flat-bottom, black 384-well plates (Corning), in a volume of 50  $\mu\text{L}$  (reaction buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl)) on a Tecan safire<sup>2</sup> plate reader. The following experimental conditions were used: 0.5 to 6.0  $\mu\text{M}$  Nedd8-AMC in reaction buffer, 4 nM CSN5<sup>AC</sup>/CSN6<sup>AC</sup> binary complex (110 nM CSN5<sup>AC</sup> and 200 nM CSN6<sup>AC</sup> were mixed prior to reaction and the concentration of the binary complex was calculated using the value of the apparent  $K_D$  value determined by ITC), 200 nM CSN5<sup>AC</sup> alone or 23 nM CSN. In the case of saturating concentrations of Nedd8-AMC, initial rate data were fitted to the Michaelis-Menten equation using least square analysis to determine  $k_{\text{cat}}$  and  $K_M$  in the Prism software (GraphPad). When saturating concentrations of Nedd8-AMC were not reached, the value of  $k_2$ , corresponding to an approximate value of the  $k_{\text{cat}}/K_M$  ratio was obtained by linear fitting of initial rate data.

### Deneddylatation assays

The CSN5<sup>AC,WT</sup>/CSN6<sup>AC</sup> complex and recombinant CSN, were diluted to 4 nM in reaction buffer (20 mM Tris-HCl pH 7.5, 50 mM NaCl) and incubated at 37°C in the presence of 1  $\mu\text{M}$  Cullin1-Nedd8/Rbx1. A time course (0 to 120 minutes) was recorded for each reaction in independent duplicates. One time point was performed for CSN5<sup>AC,R106T</sup> (900 nM) for 2 hours incubation time. The reactions were stopped in SDS sample buffer and analysed on Tris-tricine gel stained in Instant Blue solution (Expedeon, UK). Bands of Cullin-Nedd8/Rbx1 were detected and quantified using Carestream Molecular Imaging instrument connected to the Gel Logic 2200PRO software (Equilab, White-stone, NY).

### Pro-Nedd8 processing assays

This assay conditions were as follows: reaction buffer (20 mM Tris-HCl pH 7.5, 50 mM NaCl); 1  $\mu\text{M}$  pro-Nedd8 WT or variants; CSN5<sup>AC</sup> (WT or variants), alone or the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> complex was diluted to 900 nM (for CSN5<sup>AC</sup> alone) or to 4 nM (for the dimer: 110 nM CSN5<sup>AC</sup> and 200 nM CSN6<sup>AC</sup>). A time course at 37°C was performed in independent duplicates for each reaction. Samples were analysed as described for deneddylatation assays.

### Crystallization, data collection and structure determination

The coordinate and structure factor files have been deposited in the PDB under the code 4QFT. Purified CSN6<sup>AC</sup> at 8 mg mL<sup>-1</sup> was subjected to crystallisation trials using commercial screening kits. Crystals grown at 18°C, in 0.2 M Tri-Sodium citrate, 0.1 M Bis-Tris propane pH 6.5 and 20% PEG3350, were harvested, cryo-protected in the crystallisation solution supplemented with 10% glycerol and flash-frozen in liquid nitrogen. A dataset was collected at a 1.8 Å resolution at ESRF ID29 beamline and processed using MOSFLM and SCALA from CCP4 suite [25]. The structure was solved by molecular replacement using Phaser with the chain A of human Rpn8<sup>AC</sup> (residues 1–187) as the search

model (PDB code 2O95) [26]. The use of human Rpn8 fragment structure as the search model for molecular replacement is supported by the fact that (i) the boundaries of the human CSN6<sup>AC</sup>, on which we worked, better correspond to those of human Rpn8 than to those of *Drosophila* CSN6<sup>AC</sup>; (ii) the structure of *Drosophila melanogaster* CSN6<sup>AC</sup> that lacks the C-terminal portion of the MPN domain displays an atypical strand orientation at its C-terminus; (iii) the crystal structure of the human Rpn8<sup>AC</sup> template model is at the highest resolution (1.95 Å). The initial model was built by ARP/wARP into the electron density map [27]. Further refinement procedure was carried out by alternate cycles of manual rebuilding and REFMAC refinement. Data collection and refinement statistics are compiled in Table S1.

### NMR

Uniformly <sup>15</sup>N-labelled CSN6<sup>AC</sup> protein samples were dissolved in 50 mM Tris buffer, 150 mM NaCl (pH adjusted to 6.8) in 90% H<sub>2</sub>O/10% D<sub>2</sub>O. In all experiments, the <sup>1</sup>H carrier was centred on the water resonance and a WATERGATE sequence [28,29] was incorporated to suppress solvent resonances. CSN6<sup>AC</sup> assignment. Spectra were acquired at 298K and pH 6.8, with a 600  $\mu\text{M}$  CSN6<sup>AC</sup> protein sample on a 700 MHz Avance Bruker spectrometer equipped with triple-resonance (<sup>1</sup>H, <sup>15</sup>N, <sup>13</sup>C) z-gradient cryo-probe (Figure S1A). NMR data were processed using GIFA (v 4.0) [30], and Topspin (v. 2.1). The 3D <sup>1</sup>H-<sup>15</sup>N-NOESY-heteronuclear single quantum coherence spectroscopy (HSQC) spectrum was analysed using strip-plots with manual reordering of the sequential stretches according to the main chain assignment strategy [31,32,33]. Side chain assignments were carried out using the <sup>1</sup>H-<sup>15</sup>N-TOCSY-HSQC spectrum. The first two residues Gly-Pro in the N-terminal tag were not assigned. For the remaining of the sequence, assignments were obtained for <sup>15</sup>N, HN and Ha atoms (excluding Pro residues, and A71, L89, D101 and H160 residues). The overall assignment completeness for <sup>1</sup>HN-<sup>15</sup>N resonances was 98% (Tables S2, S3). CSN5<sup>AC</sup>/CSN6<sup>AC</sup> complex. An HSQC spectrum at 308 K was acquired starting from a <sup>15</sup>N-labelled CSN6<sup>AC</sup> sample at 100  $\mu\text{M}$  (Figure S1B). After addition of 200  $\mu\text{M}$  unlabelled CSN5<sup>AC</sup> to the <sup>15</sup>N-labelled CSN6<sup>AC</sup> sample, another HSQC was recorded (Figure S1B). The assignments of CSN6<sup>AC</sup> at 308K were derived from those at 298K, by recording an intermediate CSN6<sup>AC</sup> HSQC spectrum at 305K.

### Protein-protein docking and modelling

CSN5<sup>AC</sup> and CSN6<sup>AC</sup> models were treated as rigid bodies and all six rotational and translational parameters were fully explored in the *ab initio* docking program ZDOCK using 6° sampling on Euler angles. A list of 30 'passive' solvent-accessible residues of CSN6<sup>AC</sup> excluded from the interface with CSN5<sup>AC</sup> derived from the analysis of non-shifted CSN6<sup>AC</sup> peaks in HSQC experiments in the presence of CSN5<sup>AC</sup> was incorporated in the filtering option of the ZDOCK procedure (Figure S1B; Table S4). Among the 54,000 models generated, the top 2,000 complexes, ranked according to ZDOCK were further analyzed. The starting structures for human CSN5<sup>AC</sup> and CSN6<sup>AC</sup> proteins were the crystal structures solved at 2.6 Å and 1.8 Å resolution, respectively. Two docking pairs of protein models were used for docking simulations: (i) the CSN5<sup>AC</sup> (chain A, PDB code: 4F7O), with Ins-1 segment (residues 98–132) fitted to the corresponding ubiquitin-liganded AMSH-LP Ins-1 structure (AMSH-LP structure corresponds to the MPN domain: residues 264–436; PDB code 2ZNV) using MODELLER9 program, followed by manual fitting (to avoid steric clash with the C-terminal LRGG residues of ubiquitin

from AMSH-LP distal ubiquitin), combined with the 38–191 CSN6 fragment ( $\alpha 4$  deletion; referred to as CSN6<sup>AC, $\Delta\alpha 4$</sup> ); and (ii) the same CSN5<sup>AC</sup> model with the N- and C-terminal appending segments removed (residues 2–31 and 232–257; referred to as CSN5<sup>AC, $\Delta 2-31,\Delta 232-257$</sup> ) combined with CSN6<sup>AC, $\Delta\alpha 4$</sup>  (residues 38–207). We subsequently filtered and re-ranked the 2000 hits using a dedicated procedure incorporating additional experimental constraints in order to identify the correct docking conformations. These constraints were derived from mutagenesis and biochemical data (CSN6 H44, V115 residues and CSN5 E115, Y116, Y120 residues were contributing to the CSN5/CSN6 interface). Additional information was also incorporated in this filtering procedure: primary sequence conservation data for both CSN5 and CSN6 calculated using the PAT server (sequence conservation calculated for each residue in the aligned sequences, by pairwise comparisons using a BLOSUM62 matrix-derived substitution probability score and averaged/normalized for each position to have a 100% value for absolute conservation) [34], total buried area in the dimer interface (calculated using the NACCESS program (<http://www.bioinf.manchester.ac.uk/naccess>), surface shape complementarity (obtained by SC program [25,35]) and interface parameters derived from the PISA program [36]. Docking simulations were performed on Intel Linux cluster platforms. The low frequency normal modes of CSN5<sup>AC</sup> were computed with WEBnm@ server [37], providing information on regions with slow collective motions. The six first non-trivial modes were analyzed.

#### CX-MS

Roughly 100  $\mu$ l of 1.2 mg/ml CSN sample were cross-linked at 1 mM disuccinimidyl suberate d0/d12 (DSS, Creative Molecules Inc.), followed by tryptic digestion and enrichment for cross-linked peptides, essentially as described [38]. LC-MS/MS analysis was carried out on an Orbitrap Elite mass spectrometer (Thermo Electron, San Jose, CA) and data were searched using *xQuest* [39] in iontag mode with a precursor mass tolerance of 10 ppm. For matching of fragment ions tolerances of 0.2 Da for common-ions and 0.3 Da for cross-link ions were used. False discovery rates (FDR) of cross-linked peptides were assigned using *xProphet* (version 2.5.2) [40]. Cross-linked peptides were identified with a delta Score <0.95 and a linear discriminant (ld) score >25 and additionally analyzed by visual inspection in order to ensure good matches of ion series on both cross-linked peptide chains for the most abundant peaks.

#### Revisiting the topology of CSN

The starting point of the modelling work was the assembly model generated by [8]. To the individual subunit models of the initial CSN model, were substituted updated models of the following subunits generated by Phyre [41]: CSN1 (modelled on a partial crystal structure of the *Arabidopsis thaliana* (At) protein; PDB code 4LCT; [42]); CSN5 (modelled on the CSN5<sup>AC</sup> crystal structure; PDB code 4F7O; [19]); CSN6 (modelled on the CSN6<sup>AC</sup> crystal structure; this work); CSN7 (modelled on a partial crystal structure of the At protein; PDB code 3CHM; [43]); CSN8 (modelled on the crystal structure of eIF3k; PDB code 1RZ4; [44]). First the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> dimer model was placed in the EM density map and its position further refined using the CSN5<sup>AC</sup>-Nedd8/CSN6<sup>AC</sup> ternary model. The remaining newly modelled CSN subunits were placed in the density at their respective positions, as established by [8], leaving an unoccupied portion in the electron density. Parallel with the 26S proteasome lid suggested that the C-termini portions of the CSN subunits might occupy this portion. We therefore manually placed the

modelled C-termini of the CSN subunits in this density, exploiting inter-subunit cross-links (Table S5). In the last step of modelling, a Molecular Dynamics Flexible Fitting (MDFF) approach was used to improve this initial fit and the quality of the final assembly model [45]. To do so, the various CSN subunits were first rigidly fitted into the EM map (EMD 2173) [8], using the collage program in SITUS package [46]. We used the MDFF procedure with implicit solvent and optimized strength of structural restraints and steering forces [47], together with additional C $\alpha$ -C $\alpha$  Lys distance restraints to constrain the experimental inter-subunit cross-linked lysine residues. 5000 cycles of minimization were followed by 100,000 cycles of restrained steered MD. The accuracy of the model was assessed by local correlation between the model and the EM map, and by analysis of the C $\alpha$ -C $\alpha$  distance between cross-linked Lys residues.

## Results

### The MPN domains of CSN5 and CSN6 form a stable heterodimer

Previous structural studies of the CSN suggested that CSN5 and CSN6 interact [8,48]. We therefore set out to quantify the association between the CSN MPN subunits. The association of the 1–257 and 31–211 domains of CSN5 and CSN6, amenable to soluble bacterial expression, referred to as CSN5<sup>AC</sup> and CSN6<sup>A</sup>, respectively (Figures 1A, S2) was probed by ITC experiments. The data revealed a  $K_D$  of  $6 \pm 1 \mu$ M for the MPN heterodimer (Figure 1B,C). Compatible with these findings, significant changes in the HSQC spectra comprising shifts, intensity changes and disappearance of peaks were observed when CSN5<sup>AC</sup> was added to <sup>15</sup>N-uniformly labelled CSN6<sup>AC</sup> (Figure S1B). Moreover, binding of CSN6<sup>AC</sup> to a conformationally-relaxed CSN5<sup>AC</sup> variant, in which the residue arginine 106 of the Ins-1 is substituted by a threonine (CSN5<sup>AC,R106T</sup>) [19] was probed by ITC. In this previous structural, biochemical and computational analysis of CSN5<sup>AC</sup>, the Ins-1 region was identified as an important region for the activation state of CSN5<sup>AC</sup>. More specifically, in the WT inactive form of CSN5<sup>AC</sup>, the Ins-1 appears conformationally constrained and folded back onto the zinc-binding site. The conformational relaxation of the Ins-1, allowing an activity gain of the enzyme, was achieved *in silico* by loss of a salt bridge between the arginine 106 (R106) and the aspartate 151 (D151) upon substitution of R106 by a threonine, as shown in molecular dynamics simulations. This was mirrored *in vitro* by an activation of the enzyme upon Ins-1 conformational relaxation resulting from the un-anchoring of Ins-1 from the zinc-binding site. ITC measurements indicated a comparable affinity between CSN6<sup>AC</sup> and the conformationally-relaxed CSN5<sup>AC,R106T</sup> variant, to that of CSN5<sup>AC,WT</sup>/CSN6<sup>AC</sup> (Figure 1C).

### CSN6<sup>AC</sup> confers increased Nedd8 affinity and isopeptidase activity to CSN5<sup>AC</sup>

The CSN catalytic subunit lacks enzymatic activity in a stand-alone form, due to the auto-inhibitory position of Ins-1 [17,19,48]. We hypothesized that, while CSN5 by itself has a very low affinity for Nedd8, upon activation by an unknown factor, the Ins-1 region rearranges to form part of a Nedd8 recruitment site groove. Interestingly, MPN- subunits of the 26S proteasome lid and of BRCC36-containing complexes are able to activate their otherwise inactive catalytic subunits, Rpn11/POH1 and BRCC36, respectively [20]. Having shown that the N-terminal MPN domains of CSN5 and CSN6 heterodimerize, we went on to evaluate the effect of CSN6<sup>AC</sup> on CSN5<sup>AC</sup>'s affinity for Nedd8

and its isopeptidase activity on substrates ranging from protein to a physiological complex.

First we utilized a fluorescence anisotropy assay to assess the binding affinity of Alexa Fluor 488-labelled Nedd8 for CSN5<sup>AC</sup> variants (Figure S3). Consistent with our hypothesis, CSN5<sup>AC,WT</sup> alone displays the weakest affinity for Nedd8 ( $K_D = 320 \pm 59 \mu\text{M}$ ), but conformational relaxation of the Ins-1 region in CSN5<sup>AC,R106T</sup> increases the affinity to an apparent  $K_D$  of  $202 \pm 25 \mu\text{M}$ . Importantly, the CSN5<sup>AC,WT</sup>/CSN6<sup>AC</sup> (or CSN5<sup>AC,R106T</sup>/CSN6<sup>AC</sup>) complex has a similar affinity for Nedd8, with an apparent  $K_D$  of  $179 \pm 17 \mu\text{M}$  (or  $144 \pm 22 \mu\text{M}$ ). Notably, CSN6<sup>AC</sup> alone does not appear to greatly contribute to Nedd8 binding (Figure S3A).

Next the catalytic properties of the enzymes were investigated on Nedd8-AMC (Figure 1D). Consistent with previous data [19], 200 nM CSN5<sup>AC,R106T</sup> hydrolysed Nedd8-AMC at a detectable rate, whilst CSN5<sup>AC,WT</sup> showed no activity. As saturating concentrations of the substrate were not reached, we used the available titration data of CSN5<sup>AC,R106T</sup> to determine a  $k_{\text{app}}$  value of  $57 \pm 6 \text{ M}^{-1}\text{s}^{-1}$  from the linear fitting of the initial rate data. The CSN5<sup>AC,WT</sup>/CSN6<sup>AC</sup> binary complex at a 4 nM concentration, obtained by mixing 110 nM CSN5<sup>AC,WT</sup> and 200 nM CSN6<sup>AC</sup>, displayed robust catalytic efficiency ( $k_{\text{app}}$  value of  $59 \times 10^3 \pm 1 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$ ), although the activity of CSN5<sup>AC,WT</sup>/CSN6<sup>AC</sup> complex was not saturable at 6  $\mu\text{M}$  substrate concentration and therefore determination of the  $k_{\text{cat}}/K_M$  ratio was not possible (Figure 1D). For comparison with the data on the binary complex, we measured the ability of the holo-CSN to hydrolyse Nedd8-AMC, which showed saturation by the substrate (Figure 1D). CSN possesses an estimated  $k_{\text{cat}}/K_M$  ratio value of  $17 \times 10^3 \pm 1 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$  towards Nedd8-AMC.

We subsequently studied the hydrolytic activity of CSN and of the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> dimer on the physiological substrate Cullin1-Nedd8/Rbx1. The holo-CSN was very efficient at hydrolysing the isopeptide bond between Nedd8 and Cullin1, as reported in [49] (Figure 1E,F). Within less than 10 minutes, most of the Cullin1-Nedd8 substrate at 1  $\mu\text{M}$ , was hydrolysed by 4 nM CSN, as determined by gel shift assay (Figure 1E). Comparatively, the activity of 4 nM CSN5<sup>AC,WT</sup>/CSN6<sup>AC</sup> dimer on 1  $\mu\text{M}$  Cullin1-Nedd8 is lower, with a modest proportion of the substrate processed in the same period (Figure 1E,F). For context, the processing of 1  $\mu\text{M}$  Cullin1-Nedd8 by 900 nM CSN5<sup>AC,R106T</sup> alone (i.e. over 200-fold molar excess of enzyme compared to the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> conditions) after 120 min incubation revealed that the R106T variant alone has a reduced capacity to hydrolyse Cullin1-Nedd8 compared to the binary CSN5<sup>AC,WT</sup>/CSN6<sup>AC</sup> complex.

These findings may imply that the heterodimer constituted by the CSN5/CSN6 MPN domains in the stand-alone state possibly correspond to the minimal catalytic centre, significantly contributing to the isopeptidase activity of human CSN, but with low efficiency on a CSN physiological substrate. As CSN displays significantly higher deneddylase activity than the CSN5/CSN6 MPN complex on a physiological cullin substrate, this could suggest that the optimal activity of the CSN5/CSN6 MPN domains would only be revealed in the context of the holo-CSN complex. These data suggests a hierarchy in the catalytic activity over the enzymatic system and the substrate type, with the inactive CSN5<sup>AC,WT</sup> form, CSN5<sup>AC,R106T</sup> that has basal activity, the MPN heterodimer alone that has robust isopeptidase activity on non-physiological substrates and the CSN5/CSN6 heterodimer in the context of CSN that recapitulates strong activity on both non-physiological and physiological substrates.

### CSN6<sup>AC</sup> crystal structure resembles that of Rpn8<sup>AC</sup>

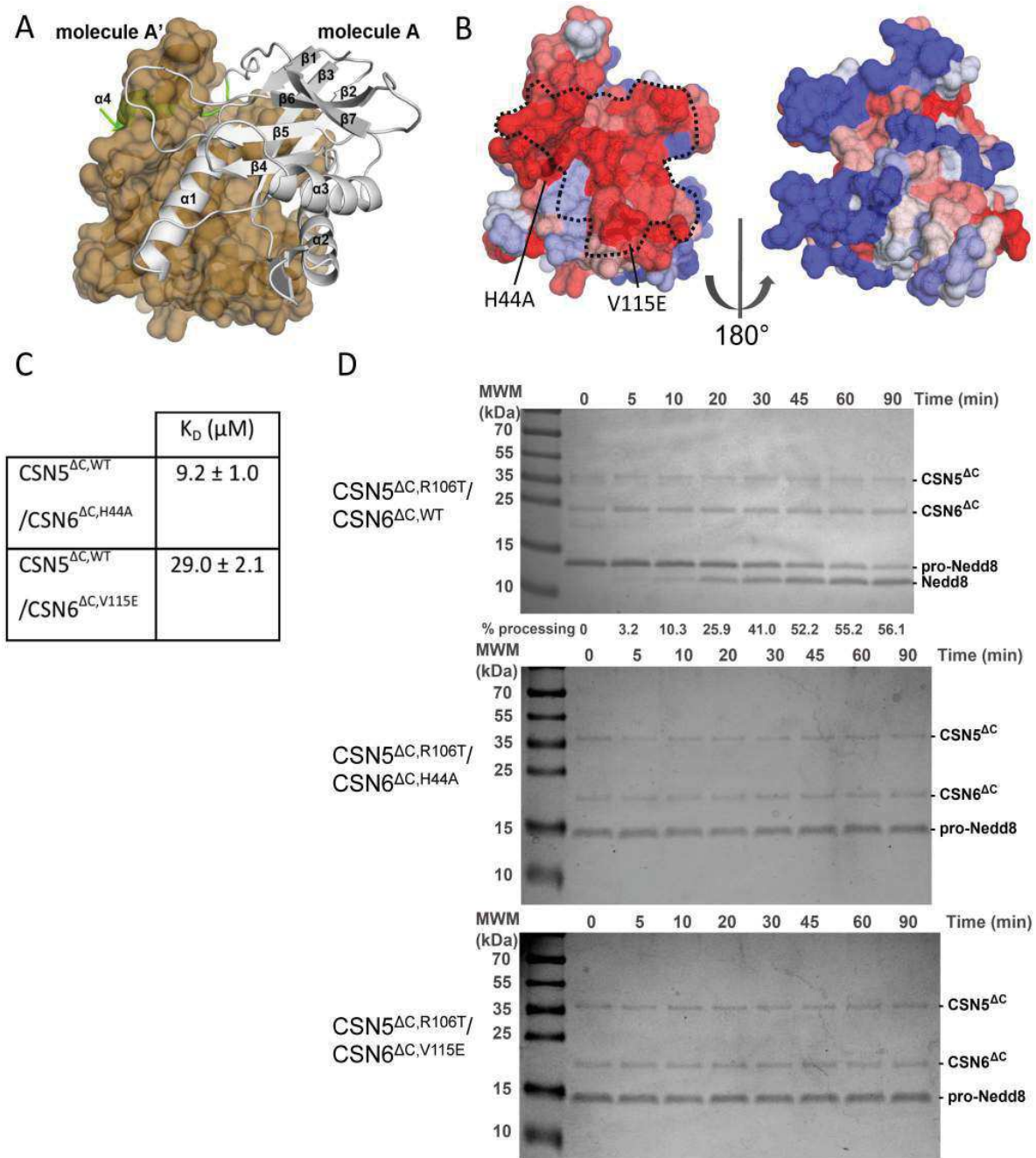
To gain additional insights into the molecular mechanism of CSN5 activation by CSN6, the crystal structure of the MPN<sup>+</sup> core fragment was determined by molecular replacement at 1.76 Å resolution, using the human Rpn8<sup>AC</sup> orthologue as the search model (PDB code 2O95; Table S1). The CSN6<sup>AC</sup> structure encompasses the typical MPN domain fold composed of a central  $\beta$ -sheet made of seven strands decorated by three  $\alpha$ -helices and a small anti-parallel  $\beta$ -sheet in the Ins-1 region. The MPN core is extended by a C-terminal appendix (residues 192 to 208) that includes a small helical portion ( $\alpha$  201 to 207; Figure 2A). The asymmetric unit contains one CSN6<sup>AC</sup> molecule, but the crystallographic 2-fold symmetry operator generates an  $\alpha$ -helix swapping-stabilised dimer, involving the C-terminal  $\alpha$ 4 helices (Figure 2A). The MPN domain dimerisation is structurally well documented, but, until recently, all the MPN dimers identified were mediated via different interfaces [14]. Interestingly, despite their sequence conservation (Figure S4A), the homodimeric organisation observed in human CSN6<sup>AC</sup> is different from that in a CSN6 51–187 fragment from *Drosophila melanogaster* [16], but identical to the human Rpn8<sup>AC</sup> and to the budding yeast Rpn11<sup>AC</sup>/Rpn8<sup>AC</sup> heterodimer crystal structures, comprising the Rpn8 2–178 or 1–176 and Rpn11 2–239 or 1–220 fragments [21,22,50] (Figure S4 B–G). Human CSN6<sup>AC</sup> dimer showed high structural homology with swapped Rpn8<sup>AC</sup> dimer, as illustrated in Figure S4B,C,F,G. 304 equivalent C $\alpha$  atoms were superimposed with a rmsd of 1.69 Å and a sequence identity of 23% using LSQMAN program [51]. The same comparison of the monomeric form of human CSN6<sup>AC</sup> with human Rpn8<sup>AC</sup> and *Drosophila* CSN6<sup>AC</sup> gave rmsd values of 1.64 Å for 153 C $\alpha$  pairs and 1.35 Å for 84 C $\alpha$  pairs, respectively.

### A highly conserved surface of CSN6<sup>AC</sup> could mediate the association with CSN5<sup>AC</sup>

Crystallisation of the human CSN5<sup>AC</sup>/CSN6<sup>AC</sup> heterodimer has so far been unsuccessful, in spite of extensive trials. Therefore, to characterise further this complex, we probed the association of CSN5<sup>AC</sup> and CSN6<sup>AC</sup> by NMR to help define the regions responsible for the interaction. Because of the difficulties to unambiguously ascribe CSN6<sup>AC</sup> residues affected by CSN5<sup>AC</sup> (owing to size and poor solubility of the binary complex), instead we extracted, through <sup>15</sup>N-CSN6<sup>AC</sup> HSQC experiments, the residues that were *not* perturbed upon addition of CSN5<sup>AC</sup> (20% of the residues; Table S4). Although not unambiguous, these NMR restraints were useful in combination with sequence conservation pattern and crystal packing information to delineate a CSN6<sup>AC</sup> surface potentially impacted by the addition of CSN5<sup>AC</sup>. The CSN6<sup>AC</sup> surface engaged in the binding to CSN5<sup>AC</sup> most likely corresponds to the region comprised between the  $\alpha$ 1/ $\alpha$ 2 helices, as delineated in Figure 2B with dotted grey line. Indeed the functional significance of this particular surface is strengthened by the facts that: (i) it is a highly conserved surface among CSN6 family members (Figures 2B,S4A) and (ii) it corresponds to the homodimer interface found in human CSN6<sup>AC</sup> and Rpn8<sup>AC</sup> crystal structures, as well as to the heterodimer interface found in budding yeast Rpn11<sup>AC</sup>/Rpn8<sup>AC</sup> (Figures 2A,S4B,C,F,G) [21,22].

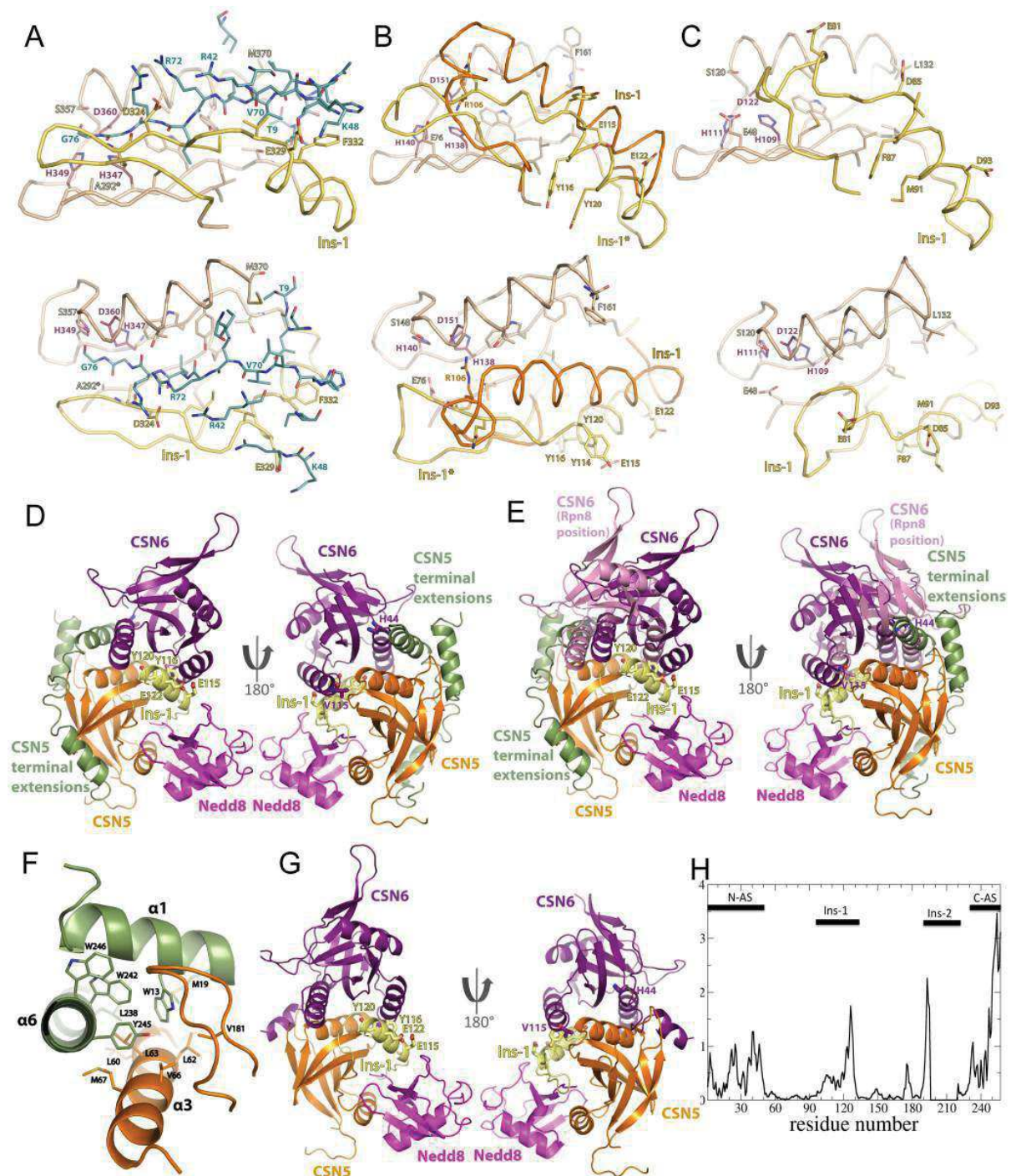
To further validate this interaction surface, we mutated two CSN6<sup>AC</sup> residues, H44 to an alanine and V115 to a glutamic acid, chosen for their high conservation and their positions on the putative interaction surface, and determined the effect of these mutations on the association with CSN5<sup>AC</sup> by ITC and activity measurements (Figures S4A,2B,C,D). The H44A variant showed a slight reduction in heterodimerisation compared with the WT





**Figure 2. CSN6<sup>ΔC</sup> binds to CSN5<sup>ΔC</sup>.** (A) Human CSN6<sup>ΔC</sup> displays a classical MPN fold. The crystallised fragment of human CSN6 contains an MPN core (white ribbon), completed by a C-terminal extension (green ribbon). The asymmetric unit contains one molecule (A); the 2-fold crystallographic symmetry-related one, A' is shown in beige as a surface. (B) CSN6<sup>ΔC</sup> binding to CSN5<sup>ΔC</sup> engages a conserved surface. CSN6 sequence conservation is displayed on the molecular surface of CSN6<sup>ΔC</sup> with a colour coding blue to red from variable to conserved, respectively, as calculated with the PAT server [34]. The dotted line corresponds to the surface fingerprint of CSN5<sup>ΔC,Δ2-31,Δ232-257</sup> in the docking model of CSN5<sup>ΔC</sup>/CSN6<sup>ΔC</sup>. (C)  $K_D$  values of CSN5<sup>ΔC</sup>/CSN6<sup>ΔC</sup> variant pairs obtained from ITC data. (D) Pro-Nedd8 processing by CSN5<sup>ΔC</sup>. Time course assay of pro-Nedd8 (1  $\mu$ M) processing by CSN5<sup>ΔC,R106T</sup>/CSN6<sup>ΔC</sup> (top; 4 nM, obtained from mixing 150 nM CSN5<sup>ΔC,R106T</sup> and 200 nM CSN6<sup>ΔC</sup>), CSN5<sup>ΔC,R106T</sup>/CSN6<sup>ΔC,H44A</sup> (middle; 4 nM) CSN5<sup>ΔC,R106T</sup>/CSN6<sup>ΔC,V115E</sup> (bottom; 4 nM), were analysed on Coomassie stained SDS-PAGE gels. The time course assay was analysed on Coomassie stained 15% Tris-tricine SDS-PAGE. Quantification of pro-Nedd8 hydrolysis is specified, when possible, at the bottom of the gel. Quantification of the pro-Nedd8 processing was carried out as detailed in the Material and Methods section. Although CSN5<sup>ΔC,WT</sup> in the presence of CSN6<sup>ΔC</sup> displays robust activity on a range of substrates, as illustrated in Figure 1D,E,F, its level remains lower to that of CSN5<sup>ΔC,R106T</sup> in the presence of CSN6<sup>ΔC</sup>. For the pro-Nedd8 gel shift assay, as illustrated in this figure, for detection purpose, it was advantageous to use the best enzymatic system available.

doi:10.1371/journal.pone.0105688.g002



**Figure 3. JAMM/MPN<sup>+</sup>-MPN<sup>-</sup> assembly. (A,B,C) Ins-1 conformation and UBL binding.** Residues discussed in the text are shown in stick mode and the zinc-coordinating residues in purple. Bottom panel: 90° rotated view with respect to the top panel. **(A) AMSH-LP MPN Ins-1 bound to distal ubiquitin.** The binding site of AMSH-LP MPN domain to distal ubiquitin (dark blue), as revealed in [18], is represented in beige and in yellow, for the Ins-1 region. The position 292, noted A292\*, in AMSH-LP MPN has been substituted to an alanine. **(B) CSN5<sup>ΔC</sup> Ins-1 region in the stand-alone form (orange) and modelled on the basis of AMSH-LP MPN Ins-1 conformation (noted Ins-1\*, yellow).** The zone corresponding to the distal ubiquitin binding site in AMSH-LP MPN (in A) is in beige. **(C) Rpn11<sup>ΔC</sup> Ins-1 region in the context of the Rpn11<sup>ΔC</sup>/Rpn8<sup>ΔC</sup> complex.** The Rpn11<sup>ΔC</sup> Ins-1 region, together with the zone corresponding to the distal ubiquitin binding site in AMSH-LP MPN (in A) are presented in yellow and beige, respectively. **(D,E,F,G) The structures, in ribbon mode, are, unless specified otherwise, coloured as follows:** CSN5<sup>ΔC</sup> or CSN5<sup>ΔC,Δ2-31,Δ232-257</sup> in orange (N- and C-terminal extensions in green; Ins-1 region in yellow), CSN6<sup>ΔC,Δ24</sup> in violet, Nedd8 in magenta. When visible,



the CSN6<sup>AC</sup> H44 and V115 residues and CSN5<sup>AC</sup> E115, Y116 and Y120 are shown in stick mode. **(D) A CSN5/CSN6 model obtained by data-driven docking.** The best CSN5<sup>AC</sup>-Nedd8/CSN6<sup>AC,Δ24</sup> pose obtained from the docking procedure is represented. Right panel: 180° rotated view of left Panel. **(E) Topology of CSN5<sup>AC</sup>/CSN6<sup>AC</sup> docking model and Rpn11<sup>AC</sup>/Rpn8<sup>AC</sup>-like CSN5<sup>AC</sup>/CSN6<sup>AC</sup>.** The best CSN5<sup>AC</sup>-Nedd8/CSN6<sup>AC,Δ24</sup> pose obtained from filtering ZDOCK solutions is represented, as in (D). To this model, the CSN5<sup>AC</sup>-Nedd8/CSN6<sup>AC,Δ24</sup> model built following the Rpn11<sup>AC</sup>/Rpn8<sup>AC</sup> structure has been superimposed. The CSN6<sup>AC,Δ24</sup> structure that is superimposed onto Rpn8<sup>AC</sup> is shown in pink to distinguish it from the CSN6<sup>AC,Δ24</sup> obtained in our docking solution. Right panel: 180° rotated view of left Panel. **(F) The N- and C-terminal helices specific to CSN5<sup>AC</sup>.** The N- and C-terminal helical extensions of CSN5<sup>AC</sup> are stacked onto the α3 helix and the loop linking the β6–β7 anti-parallel sheet. The residues involved in packing are shown in stick mode. **(G) CSN5<sup>AC,Δ2–31,Δ232–257</sup>/CSN6<sup>AC,Δ24</sup> docking pose.** The best docking pose using a CSN5<sup>AC</sup> model in which the N- and C-terminal extensions were deleted (CSN5<sup>AC,Δ2–31,Δ232–257</sup>) is shown. Right panel: 180° rotated view of left Panel. **(H) The normalized squared displacement of the CSN5<sup>AC</sup> Cu atoms in the lowest frequency mode 7 from normal mode analysis.** doi:10.1371/journal.pone.0105688.g003

control, whereas the V115E variant showed a significant 5-fold reduction in binding (Figures 2C, 1C). It is noteworthy that, although the interaction between CSN5<sup>AC</sup> and CSN6<sup>AC</sup> was impaired by the H44A or to a stronger extent by the V115A substitution, the effect of these variants remains modest (2 to 5-fold). The interface mediating the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> complex obtained from the molecular docking encompasses 1,430 to 1,695 Å<sup>2</sup>, depending on the type of heterodimer chosen. Prediction of 'hot spot' residues making a dominant contribution to the binding free energy is a difficult task. Here the buried surface is sizeable and could resist a single substitution not targeting a 'hot spot', whilst affecting a topological modification that would decrease the activation of CSN5<sup>AC</sup> by the mutated CSN6<sup>AC</sup> form(s). We subsequently probed the hydrolytic activity of the CSN5<sup>AC, R106T</sup>/CSN6<sup>AC, H44A or V115E</sup> complex on the Nedd8 precursor, pro-Nedd8. Indeed, our *in vitro* work suggested that the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> complex is able to process pro-Nedd8. Although CSN5 is an unlikely factor for pro-Nedd8 processing [52], we have exploited it as a reaction tool to substitute for Nedd8-AMC synthetic substrate, in a gel shift assay, where pro-Nedd8 processing was followed by gel quantification of the appearing Nedd8 band (Figure 2D). Using this pro-Nedd8 gel shift assay, we showed that the hydrolytic activity of the CSN5<sup>AC, R106T</sup>/CSN6<sup>AC, H44A or V115E</sup> variant complexes was entirely compromised (Figure 2D).

Taken together and in the context of sequence conservation and of NMR HSQC data, these results suggest an implication of the residues H44 and V115 in CSN5<sup>AC</sup>/CSN6<sup>AC</sup> binding - although these residues are possibly not located in an interfacial 'hotspot' - and in CSN6<sup>AC</sup>-mediated activation of CSN5<sup>AC</sup>, reinforcing the view that the CSN5/CSN6 MPN heterodimer is likely to involve a highly conserved surface exploited in other MPN dimers [21,22,50]. Further structural confirmation of the involvement of this surface would be required (see Notes from Authors).

### Proposed model of the CSN5/CSN6 MPN domain heterodimer

Having identified a plausible CSN5<sup>AC</sup>/CSN6<sup>AC</sup> interaction interface, we attempted to further explore the assembly of the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> dimer, molecular docking of the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> complex guided by experimental constraints. It is now well established that the incorporation of a few biochemical and NMR-derived constraints in docking methods suffices to determine the structure of a complex to high precision [53].

To help the docking procedure, a complex composed of CSN5<sup>AC</sup>, harbouring an AMSH-LP-like Ins-1 conformation and of Nedd8, was assembled. To do so, we probed the binding of Nedd8 on CSN5<sup>AC</sup> and reciprocally, to investigate whether Nedd8 binds to CSN5<sup>AC</sup> in a mode reminiscent to that of the distal ubiquitin to the MPN domain of AMSH-LP (or AMSH). The choice of the interface residues to be analysed was guided by the structurally characterised AMSH-LP MPN domain/distal ubiquitin association (Figure 3A,B) [18]. Among the two CSN5<sup>AC</sup> Ins-1

glutamate positions probed (E115, E122), the E122D variant kept strong processing activity on pro-Nedd8, whilst the E115D one displayed weak activity (Fig S5A). The importance of F332 in AMSH-LP for the distal ubiquitin binding prompted us to evaluate the role of two aromatic residues in CSN5<sup>AC</sup> Ins-1 region (Y116, Y120) on the hydrolysis of the Nedd8 precursor (Fig S5A). Additionally, effects of four pro-Nedd8 variants (T9M, R42D, K48A, V70D) on processing activity were assessed (Fig S5B). Activity measurement data on the chosen CSN5<sup>AC</sup> and pro-Nedd8 variants showed that the CSN5<sup>AC, R106T, E115A, Y116D or Y120D</sup> variants have compromised pro-Nedd8 processing activity and that the pro-Nedd8 ones, namely pro-Nedd8<sup>T9M, R42D, K48A or V70D</sup>, are important for hydrolysis by CSN5<sup>AC</sup>, validating that Nedd8 on CSN5<sup>AC</sup> probably orient similarly to the distal ubiquitin and AMSH-LP MPN (and Sst2, the AMSH homolog in yeast) (Figures 3A,B, S5A,B) [18,54].

To assemble the CSN5<sup>AC</sup>-Nedd8 complex we then modelled the CSN5 Ins-1 segment to resemble that of AMSH-LP MPN. Interestingly, as a result of the modelling procedure, CSN5<sup>AC</sup> modelled Ins-1 resembles the conformation of Rpn11<sup>AC</sup> Ins-1 in the binary complex with Rpn8<sup>AC</sup> (Figure 3C). Exploiting the above detailed finding that CSN5<sup>AC</sup> and Nedd8 interaction proceeds through similar surfaces to that of AMSH-LP-distal ubiquitin's (Figures 3A,B, S5A,B), a complex composed of CSN5<sup>AC</sup> displaying an AMSH-LP-like Ins-1 conformation and of Nedd8 was assembled and used to probe the association of CSN5<sup>AC</sup> and CSN6<sup>AC</sup>.

Subsequently, molecular docking simulations of the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> dimer were performed by the software ZDOCK followed by a tailored post-docking filtering strategy. The simulations resulted in a cluster of similar docking poses with favourable interfacial metrics and satisfying all our experimental data (Figures 3D, S6). In the resulting model, CSN5<sup>AC</sup> and CSN6<sup>AC</sup> are placed in a topology that happens to be close to that observed in the crystallographic human Rpn8<sup>AC</sup> and human CSN6<sup>AC</sup> homodimers (Figures 3D, S4B,C,F,G) [50].

The surfaces engaged in the CSN5<sup>AC</sup>/CSN6<sup>AC,Δ24</sup> model are highly similar to those of the Rpn11<sup>AC</sup>/Rpn8<sup>AC</sup> complex (Figures 3E, S6). However, superimposition of their MPN<sup>+</sup>/JAMM subunits revealed that the CSN6<sup>AC,Δ24</sup> position on CSN5<sup>AC</sup> surface in our model is rotated and translated, relative to the Rpn8<sup>AC</sup> position on Rpn11<sup>AC</sup> (Figure 3E). This difference is the result of the presence of the CSN5-specific N- and C-terminal helical appendices on each side of the MPN domain (residues 6 to 17 and 235 to 251, respectively). The CSN5 specific N-terminal extension wraps around its MPN core and C-terminal helix (Figure 3F) [19]. For CSN5<sup>AC</sup> and CSN6<sup>AC</sup> to arrange in a Rpn11<sup>AC</sup>/Rpn8<sup>AC</sup>-like dimer, it would take the CSN5-specific N- and C-terminal α-helices to unstaple from its MPN core. To confirm this, docking carried out on a CSN5<sup>AC</sup> variant lacking the N- and C-terminal helices (fragment 32–232) yielded top scoring solutions corresponding to a pose of the Rpn11<sup>AC</sup>/Rpn8<sup>AC</sup> topology (Figures 3G, S6). These helices were shown to be

conformationally malleable by normal mode analysis [55] of CSN5<sup>AC</sup>, where a subset of low frequency modes shows their propensity to undergo large-scale shear and hinge motions (Figures 3H, S7). These secondary structure elements, stabilised by numerous weak interactions, could be unhooked from the core domain (Figure 3F). This analysis could strengthen the view that CSN5<sup>AC</sup> and CSN6<sup>AC</sup> interact through an architecture reminiscent of the Rpn11<sup>AC</sup>/Rpn8<sup>AC</sup> complex.

### The molecular architecture of the intact CSN

Finally, we wanted to interpret our CSN5<sup>AC</sup>/CSN6<sup>AC</sup> model in the context of the full CSN complex. We adopted a hybrid structural approach, which involved CX-MS of the intact complex [38], cross-link-guided docking of individual subunits and domains and molecular dynamics flexible fitting of the obtained model using MDFF procedure in the CSN density of the previously published CSN-SCF EM map (EMD 2173; Table S5; Figure S8) [8,47]. For the docking we used the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> model generated in this study and for the PCI domains and missing portions of the remaining subunits, we predicted structures using the server Phyre [41] (Figures 4A, S8; Table S6). CSN1 and CSN7 models were largely based on their crystal structures (PDB codes 4LCT, 3CHM, respectively) [42,43]. The CSN5<sup>AC</sup>/CSN6<sup>AC</sup> model described above was fitted as a rigid body (Figure 4A) and docked in the previously assigned EM density portion [8]. However, using our ternary complex model composed of CSN5<sup>AC</sup>-Nedd8/CSN6<sup>AC</sup> (Figure 4A) for refinement, resulted in a slightly different orientation. In our model, CSN6<sup>AC</sup> moved to a more peripheral location, relative to the one previously reported [8]. This resulted in a conspicuous density segment left empty in the thus re-docked CSN model (Figure 4A, red star).

Multiple cross-links between the extreme C-terminal sequence portions of peripheral and central subunits, like CSN7-CSN1 or CSN5-CSN1 indicated that these elements contribute a major, previously unrecognised binding interface (Table S5, Figure S8). Importantly, most of these C-terminal elements are predicted to adopt helical structures (Figure S8). This led us to hypothesize that they all join in a central helical bundle in a topology reminiscent of the 26S proteasome lid [23]. To test this hypothesis we generated models for the C-termini of all CSN subunits and arranged them in a compact helical bundle like in [23] (Figures 4A,B, S8). The resulting cluster matched well the corresponding density in shape and size (Figure 4B), although the precise orientation and topology of the helices could not be optimized due to the low resolution of the EM density map. Potential positions of the C-termini portions contributing to the central bundle and of the CSN subunits were refined to best satisfy the set of inter-subunit cross-links providing powerful constraints (21 inter-subunits cross-links between the 8 C-terminus subunits; Table S5).

The local cross-correlation coefficient between the revisited CSN assembly and the EM map ( $ccc=0.84$ ) has increased compared to the initial model ( $ccc=0.80$ ). Both models for the eight subunits contained the same number of residues (~2,750), but positioned differently inside the 25 Å resolution map, in the particular case of CSN5/CSN6 and of the C-terminal segments. More significantly, the average distance between lysine Cαs of the 70 identified intra- and inter- cross-links is 32.9 Å (29.5 Å if 4 clear outlier inter-subunit cross-links with distance >60 Å are excluded from the analysis). This results in an improvement over the initial model for which the average distance of the 55 cross-links was ~38 Å, suggesting that our CX-MS guided procedure has improved the overall subunit topology within the CSN. Illustrative examples of the agreement between identified cross-links with the revisited model, specifically relevant to the helical bundle and to

the CSN5/CSN6 dimer with respect to the rest of the edifice are between the C-terminus of CSN1 and CSN3 (distance K467-K418: 23 Å); CSN6 and CSN7 (distance K199-K199: 27 Å); CSN2 and CSN5 (distance K426-K180: 22 Å); CSN5 and CSN7 (distance K180-K199: 32 Å). However it is noteworthy that some of the observed cross-links linkages are difficult to reconcile with the overall topology of the molecular model, as for example the link between CSN1 K97 and CSN2 K415 and between CSN1 K422 and CSN3 K115. Some of these cross-links, for which the distance in the model exceeds the maximal distance that one can expect between two cross-linked residues, may arise from conformational variability within parts of the CSN, as recently demonstrated for subunits Rpn5 and Rpn6 of the 26S proteasome lid, corresponding to CSN4 and CSN2, respectively [12,56]. It is therefore interesting to speculate, that the observed cross-link distribution may therefore at least partially be influenced by an intrinsically dynamic behaviour of the CSN [8,11].

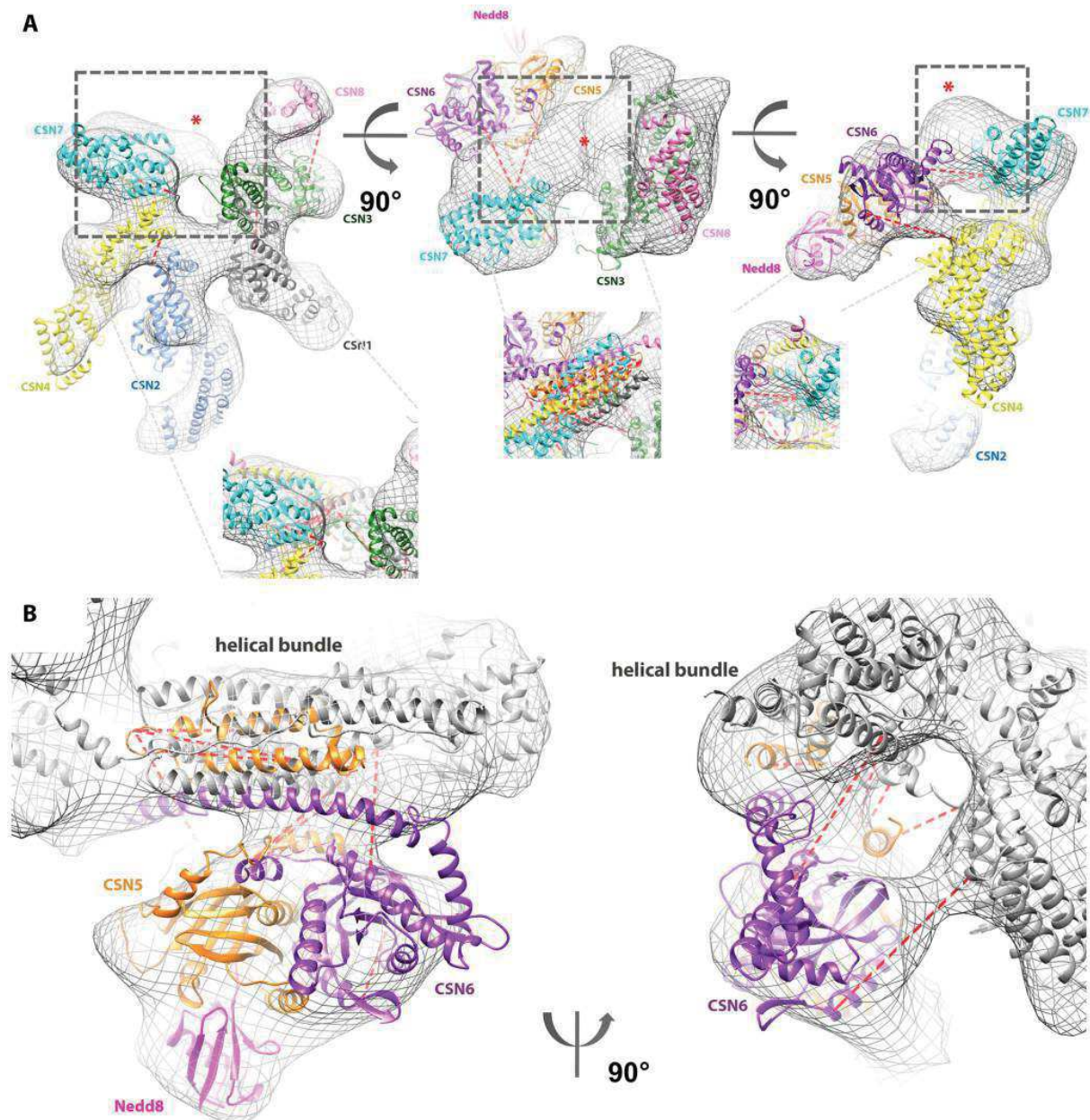
### Discussion

By dissecting biochemical and structural elements of the CSN, we have demonstrated that CSN6<sup>AC</sup> binds to the MPN domain of the CSN5 catalytic subunit and this association increases CSN5<sup>AC</sup> affinity for Nedd8 and enhances its hydrolytic activity on a variety of substrates, suggesting that the MPN heterodimer could contribute significantly to the catalytic activity of the human CSN. The activation of an MPN<sup>+</sup>/JAMM protein mediated by a MPN<sup>-</sup> subunit was described in two other MPN<sup>+</sup>/JAMM-MPN<sup>-</sup> systems, namely the 26S proteasome lid and BRCC36-containing complexes [20,21,22]. This mechanism of activity control of MPN<sup>+</sup>/JAMM enzymes would therefore appear to be a general assembly and regulatory principle of the MPN<sup>+</sup>/JAMM-MPN<sup>-</sup> containing complexes and as such would constitute an important regulatory mechanism of these complexes [14,57].

The stimulating role of CSN6<sup>AC</sup> in CSN5<sup>AC</sup> activation demonstrated here should be considered in the light of the data on the CSN reported in [58]. The authors probed the function of CSN6 and its yeast ortholog (Csi1) MPN domain in the human and yeast CSNs, respectively. These experiments, carried out using budding yeast and mammalian CSNs and neddylated cullins purified from yeast and mammalian cells, concluded that only the C-terminal part of CSN6 is necessary for the deneddylase activity. Further investigations are needed to reconcile those results with our observations.

Taken together, our fluorescence anisotropy binding experiments and the activity measurements suggest that CSN6<sup>AC</sup> does not contribute significantly to Nedd8 direct recruitment and activates CSN5<sup>AC</sup> through at least two mechanisms. First, the catalytic activity efficiency can be achieved by an increase of affinity for Nedd8 observed either in the CSN5/CSN6 complex or by the conformational relaxation of the Ins-1 segment. Secondly, the increased activity can also result from a yet unknown mechanism. Indeed, CSN5<sup>AC,WT</sup>/CSN6<sup>AC</sup> has a higher catalytic efficiency than CSN5<sup>AC,R106T</sup> alone, although their affinity for Nedd8 is of the same order of magnitude. The contribution of CSN6<sup>AC</sup> to CSN5<sup>AC</sup> activation might therefore not be confined to the enhancement of substrate binding.

Assessment of the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> complex and of CSN on different Nedd8-derived substrates provides a molecular ranking of the catalytic activity. The higher hydrolysis rate of CSN on neddylated Cullin1 than on Nedd8-AMC supports the observation that CSN is a cullin deneddylase [7], underlying the contribution of the CSN edifice in the efficient processing of neddylated CRLs. This is therefore in agreement with the observations in [8], that



**Figure 4. The CSN topology.** Inter-subunit cross-links between CSN5 or CSN6 Lys and other CSN subunits ( $C\alpha$  distances  $<40$  Å) are marked with red dotted lines. (A) **The CSN assembly revisited**, displayed in the EM density obtained by [8] (grey mesh). A red star indicates an unoccupied electron density region. Close-up views of this region are shown after fitting of the helical bundle. (B) **CSN5-Nedd8/CSN6 complex positioned in the EM envelope**. CSN5 is shown in orange, CSN6 in purple and Nedd8 in magenta. CSN subunits (other than CSN5 and CSN6) are shown in white for context.

doi:10.1371/journal.pone.0105688.g004

PCI subunits contribute to the recruitment of the CRLs. Indeed EM studies on CSN/CRL supercomplexes highlighted the direct interaction between CSN2, and probably CSN4, and the cullin subunit of the CRL. Taken together, these data show that the CSN is an intricate assembly that is tailored to function on CRLs.

To further probe CSN assembly, we performed molecular docking of the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> complex guided by experimental constraints (Figure S6). The MPN domain heterodimerisation appears to be an important aspect in the regulation of several multiprotein complexes in eukaryotes. Vexingly, different MPN-containing domains appear to utilise different dimerisation



surfaces [14], although recent structures (budding yeast Rpn11<sup>AC</sup>/Rpn8<sup>AC</sup> [21,22]; human CSN6<sup>AC</sup>) exhibit a dimerisation mode consistent with the one originally described for human Rpn8<sup>AC</sup> [50]. Interestingly our model highly resembles the crystal structure of the Rpn11<sup>AC</sup>/Rpn8<sup>AC</sup> dimer [21,22] and the association of CSN6<sup>AC</sup> and of Rpn8<sup>AC</sup> in the direct vicinity of the Ins-1 region of CSN5<sup>AC</sup> and Rpn11<sup>AC</sup>, respectively, further underlines the essential role of Ins-1 as a regulatory switch in these enzymes.

Moreover, we explored the molecular assembly of the intact CSN by reinterpreting a published EM density map, using molecular docking of our model of the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> complex and experimental data from CX-MS experiments. This work provided an updated model of the CSN structure that comprises the repositioning of the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> dimer and a central helical bundle, reminiscent of the 26S proteasome lid and suggests that this bundle may be contributing to the assembly of the CSN [23]. This structural analysis, complementing our work on the MPN domains of CSN5 and CSN6 that highlights the importance of the MPN domain interaction in the catalytic activity, could suggest that the C-terminal regions of CSN5 and CSN6 contribute to anchoring and optimal positioning these subunits within the CSN assembly.

## Supporting Information

**Figure S1 NMR HSQC spectra of <sup>15</sup>N-CSN6<sup>AC</sup>.**  
(PDF)

**Figure S2 SDS-Page of CSN5<sup>AC</sup> and CSN6<sup>AC</sup> proteins.**  
(PDF)

**Figure S3 Fluorescence anisotropy data on CSN5<sup>AC,WT</sup>, CSN5<sup>AC,R106T</sup> and CSN6<sup>AC</sup>, on CSN5<sup>AC,WT</sup>/CSN6<sup>AC</sup> and CSN5<sup>AC,R106T</sup>/CSN6<sup>AC</sup> and their associated apparent dissociation constants.**  
(PDF)

**Figure S4 CSN6 sequences and MPN domain structure.**  
(PDF)

**Figure S5 Pro-Nedd8 processing by CSN5<sup>AC</sup>.**  
(PDF)

**Figure S6 Interface residue of the CSN5<sup>AC</sup>/CSN6<sup>AC</sup> heterodimer models expressed as the number of atomic contact pairs between residues of the predicted heterodimers.**  
(PDF)

## References

- Lydeard JR, Schulman BA, Harper JW (2013) Building and remodelling Cullin-RING E3 ubiquitin ligases. *EMBO Rep* 14: 1050–1061.
- Nalepa G, Rolfe M, Harper JW (2006) Drug discovery in the ubiquitin-proteasome system. *Nat Rev Drug Discov* 5: 596–613.
- Bassermann F, Eichner R, Pagano M (2014) The ubiquitin proteasome system - implications for cell cycle control and the targeted treatment of cancer. *Biochim Biophys Acta* 1843: 150–162.
- Deshaies RJ, Joazeiro CA (2009) RING domain E3 ubiquitin ligases. *Annu Rev Biochem* 78: 399–434.
- Duda DM, Borg LA, Scott DC, Hunt HW, Hammel M, et al. (2008) Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation. *Cell* 134: 995–1006.
- Saha A, Deshaies RJ (2008) Multimodal activation of the ubiquitin ligase SCF by Nedd8 conjugation. *Mol Cell* 32: 21–31.
- Lyapina S, Cope G, Shevchenko A, Serino G, Tsuge T, et al. (2001) Promotion of NEDD-CUL1 conjugate cleavage by COP9 signalosome. *Science* 292: 1382–1385.
- Enchev RI, Scott DC, da Fonseca PC, Schreiber A, Monda JK, et al. (2012) Structural basis for a reciprocal regulation between SCF and CSN. *Cell Rep* 2: 616–627.
- Beck F, Unverdorben P, Bohn S, Schweitzer A, Pfeifer G, et al. (2012) Near-atomic resolution structural model of the yeast 26S proteasome. *Proc Natl Acad Sci U S A* 109: 14870–14875.
- da Fonseca PC, He J, Morris EP (2012) Molecular model of the human 26S proteasome. *Mol Cell* 46: 54–66.
- Enchev RI, Schreiber A, Beuron F, Morris EP (2010) Structural insights into the COP9 signalosome and its common architecture with the 26S proteasome lid and eIF3. *Structure* 18: 518–527.
- Lander GC, Estrin E, Matyskiela ME, Bashore C, Nogales E, et al. (2012) Complete subunit architecture of the proteasome regulatory particle. *Nature* 482: 186–191.
- Querol-Audi J, Sun C, Vogan JM, Smith MD, Gu Y, et al. (2013) Architecture of human translation initiation factor 3. *Structure* 21: 920–928.
- Birol M, Echalié A (2014) Structure and Function of MPN (Mpr1/Pad1 N-Terminal) Domain-Containing Proteins. *Curr Protein Pept Sci*.
- Maytal-Kivity V, Reis N, Hofmann K, Glickman MH (2002) MPN+, a putative catalytic motif found in a subset of MPN domain proteins from eukaryotes and prokaryotes, is critical for Rpn11 function. *BMC Biochem* 3: 28.

**Figure S7 CSN5<sup>AC</sup> normal modes analysis using WEBnm@ server.**  
(PDF)

**Figure S8 Linkage map of the CSN subunits obtained from cross-linked CSN followed by MS analysis.**  
(PDF)

**Table S1 X ray data collection and refinement statistics for human CSN6<sup>AC</sup> crystal.**  
(PDF)

**Table S2 NMR experiments acquired for chemical shift assignments.**  
(PDF)

**Table S3 NMR experiments acquired for CSN5<sup>AC</sup>/CSN6<sup>AC</sup> mapping.**  
(PDF)

**Table S4 Unperturbed CSN6<sup>AC</sup> residues upon addition of CSN5<sup>AC</sup> as evaluated by NMR HSQC.**  
(PDF)

**Table S5 Cross-links used in this study.**  
(PDF)

**Table S6 Sequences of the CSN subunit models used.**  
(PDF)

## Acknowledgments

The authors thank D. Xirodimas and M. Banfield for discussions and useful reagents, X. Morelli for access to the ITC machine, D. Flot for help during data collection and J. Gracy for the sequence analysis. We acknowledge the support of ?; HPC@LR a High-Performance Computing Centre, funded by the Languedoc-Roussillon, European Union and the University Montpellier 2 (UM2) for part of the docking simulations. AE is part of the ITN UPstream and COST Proteostasis consortia.

### Note from the authors

Whilst our paper was in the latest stage of revision, a manuscript describing the X-ray crystal structure of the human CSN at 3.8 Å resolution was published [59]. Findings described in this paper support the data and conclusions reached in our study on the importance of CSN6 in CSN5 activation, the CSN5/CSN6 heterodimer and the CSN architecture.

## Author Contributions

Conceived and designed the experiments: AE FS RA. Performed the experiments: MB YY RIE FS SB AP FH CD AE. Analyzed the data: MB RIE FS MP CD AE. Contributed reagents/materials/analysis tools: CD. Contributed to the writing of the manuscript: AE MB RIE RA MP CD.

16. Zhang H, Gao ZQ, Wang WJ, Liu GF, Shtykova EV, et al. (2012) The crystal structure of the MPN domain from the COP9 signalosome subunit CSN6. *FEBS Lett* 586: 1147–1153.
17. Cope GA, Suh GS, Aravind L, Schwarz SE, Zipursky SL, et al. (2002) Role of predicted metalloprotease motif of Jab1/Csn5 in cleavage of Nedd8 from Cull1. *Science* 298: 608–611.
18. Sato Y, Yoshikawa A, Yamagata A, Mimura H, Yamashita M, et al. (2008) Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains. *Nature* 455: 358–362.
19. Echaliier A, Pan Y, Birol M, Tavernier N, Pintard L, et al. (2013) Insights into the regulation of the human COP9 signalosome catalytic subunit, CSN5/Jab1. *Proc Natl Acad Sci U S A* 110: 1273–1278.
20. Patterson-Fortin J, Shao G, Bretscher H, Messick TE, Greenberg RA (2010) Differential regulation of JAMM domain deubiquitinating enzyme activity within the RAP80 complex. *J Biol Chem* 285: 30971–30981.
21. Pathare GR, Nagy I, Sledz P, Anderson DJ, Zhou HJ, et al. (2014) Crystal structure of the proteasomal deubiquitylation module Rpn8-Rpn11. *Proc Natl Acad Sci U S A*.
22. Worden EJ, Padovani C, Martin A (2014) Structure of the Rpn11-Rpn8 dimer reveals mechanisms of substrate deubiquitination during proteasomal degradation. *Nat Struct Mol Biol*.
23. Estrin E, Lopez-Blanco JR, Chacon P, Martin A (2013) Formation of an intricate helical bundle dictates the assembly of the 26S proteasome lid. *Structure* 21: 1624–1635.
24. Crow A, Hughes RK, Taieb F, Oswald E, Banfield MJ (2012) The molecular basis of ubiquitin-like protein NEDD8 deamidation by the bacterial effector protein Cif. *Proc Natl Acad Sci U S A* 109: E1830–1838.
25. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67: 235–242.
26. McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr* 63: 32–41.
27. Langer G, Cohen SX, Lamzin VS, Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3: 1171–1179.
28. Piotto M, Saudek V, Sklenar V (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR* 2: 661–665.
29. Sklenar V, Peterson RD, Rejante MR, Feigon J (1993) Two- and three-dimensional HCN experiments for correlating base and sugar resonances in <sup>15</sup>N,<sup>13</sup>C-labeled RNA oligonucleotides. *J Biomol NMR* 3: 721–727.
30. Malliavin TE, Pons JL, Delsuc MA (1998) An NMR assignment module implemented in the Gifa NMR processing program. *Bioinformatics* 14: 624–631.
31. Englander SW, Wand AJ (1987) Main-chain-directed strategy for the assignment of <sup>1</sup>H NMR spectra of proteins. *Biochemistry* 26: 5953–5958.
32. Wand AJ, Nelson SJ (1991) Refinement of the main chain directed assignment strategy for the analysis of <sup>1</sup>H NMR spectra of proteins. *Biophys J* 59: 1101–1112.
33. Roth SM, Schneider DM, Strobel LA, Van Berkum MF, Means AR, et al. (1992) Characterization of the secondary structure of calmodulin in complex with a calmodulin-binding domain peptide. *Biochemistry* 31: 1443–1451.
34. Gracy J, Chiche L (2005) PAT: a protein analysis toolkit for integrated biocomputing on the web. *Nucleic Acids Res* 33: W65–71.
35. Lawrence MC, Colman PM (1993) Shape complementarity at protein/protein interfaces. *J Mol Biol* 234: 946–950.
36. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372: 774–797.
37. Hollup SM, Salensminde G, Reuter N (2005) WEBnm@: a web application for normal mode analyses of proteins. *BMC Bioinformatics* 6: 52.
38. Leitner A, Walzthoeni T, Aebersold R (2014) Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-MS/MS and the xQuest/xProphet software pipeline. *Nat Protoc* 9: 120–137.
39. Rinner O, Seebacher J, Walzthoeni T, Mueller LN, Beck M, et al. (2008) Identification of cross-linked peptides from large sequence databases. *Nat Methods* 5: 315–318.
40. Walzthoeni T, Claassen M, Leitner A, Herzog F, Bohn S, et al. (2012) False discovery rate estimation for cross-linked peptides identified by mass spectrometry. *Nat Methods* 9: 901–903.
41. Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* 4: 363–371.
42. Lee JH, Yi L, Li J, Schweitzer K, Borgmann M, et al. (2013) Crystal structure and versatile functional roles of the COP9 signalosome subunit 1. *Proc Natl Acad Sci U S A* 110: 11845–11850.
43. Dessau M, Halimi Y, Erez T, Chomsky-Hecht O, Chamovitz DA, et al. (2008) The Arabidopsis COP9 signalosome subunit 7 is a model PCI domain protein with subdomains involved in COP9 signalosome assembly. *Plant Cell* 20: 2815–2834.
44. Wei Z, Zhang P, Zhou Z, Cheng Z, Wan M, et al. (2004) Crystal structure of human eIF3k, the first structure of eIF3 subunits. *J Biol Chem* 279: 34983–34990.
45. Vashist H, Skinotis G, Brooks CL, 3rd (2012) Using enhanced sampling and structural restraints to refine atomic structures into low-resolution electron microscopy maps. *Structure* 20: 1453–1462.
46. Wriggers W, Milligan RA, McCammon JA (1999) Situs: A package for docking crystal structures into low-resolution maps from electron microscopy. *J Struct Biol* 125: 185–195.
47. Trabuco LG, Villa E, Mitra K, Frank J, Schulten K (2008) Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* 16: 673–683.
48. Sharon M, Mao H, Boeri Erba E, Stephens E, Zheng N, et al. (2009) Symmetrical modularity of the COP9 signalosome complex suggests its multifunctionality. *Structure* 17: 31–40.
49. Emberley ED, Mosadeghi R, Deshaies RJ (2012) Deconjugation of Nedd8 from Cull1 is directly regulated by Skp1-F-box and substrate, and the COP9 signalosome inhibits deneddylated SCF by a noncatalytic mechanism. *J Biol Chem* 287: 29679–29689.
50. Sanches M, Alves BS, Zanchin NI, Guimaraes BG (2007) The crystal structure of the human Mov34 MPN domain reveals a metal-free dimer. *J Mol Biol* 370: 846–855.
51. Kleywegt GJ (1999) Experimental assessment of differences between related protein crystal structures. *Acta Crystallogr D Biol Crystallogr* 55: 1878–1884.
52. O'Donoghue JE, Bech-Otschir D, Larsen IB, Wallace M, Hartmann-Petersen R, et al. (2013) Nedd8 processing enzymes in *Schizosaccharomyces pombe*. *BMC Biochem* 14: 8.
53. Rodrigues JP, Bonvin AM (2014) Integrative computational modeling of protein interactions. *FEBS J* 281: 1988–2003.
54. Shrestha RK, Ronau JA, Davies CW, Guenette RG, Strieter ER, et al. (2014) Insights into the mechanism of deubiquitination by JAMM deubiquitinases from co-crystal structures of enzyme with substrate and product. *Biochemistry*.
55. Dobbins SE, Lesk VI, Sternberg MJ (2008) Insights into protein flexibility: The relationship between normal modes and conformational change upon protein-protein docking. *Proc Natl Acad Sci U S A* 105: 10390–10395.
56. Unverdorben P, Beck F, Sledz P, Schweitzer A, Pfeifer G, et al. (2014) Deep classification of a large cryo-EM dataset defines the conformational landscape of the 26S proteasome. *Proc Natl Acad Sci U S A* 111: 5544–5549.
57. Pick E, Hofmann K, Glickman MH (2009) PCI complexes: Beyond the proteasome, CSN, and eIF3 Troika. *Mol Cell* 35: 260–264.
58. Pick E, Golan A, Zimble JZ, Guo L, Sharaby Y, et al. (2012) The minimal deneddylase core of the COP9 signalosome excludes the Csn6 MPN- domain. *PLoS One* 7: e43980.
59. Lingaraju GW, Bunker RD, Cavadini S, Hess D, Hassiepen U, et al. (2014) Crystal structure of the human COP9 signalosome. *Nature* in press.

## **Patent filed**

### **Csn5 polypeptides and uses thereof for screening therapeutic agents**

**Publication number:** WO2014057134 A1

**Date:** Apr 17, 2014

**Summary:** The present invention relates to mutated CSN5 polypeptides and their use in a method of screening modulators of CSN5 activity that could be used as therapeutic agents.